

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»**

**Факультет інформатики та обчислювальної техніки  
Обчислювальної техніки**

До захисту допущено:

Завідувач кафедри

\_\_\_\_\_Сергій СТИПЕНКО

«\_\_\_»\_\_\_\_\_20\_\_р.

**Дипломний проєкт**

**на здобуття ступеня бакалавра**

**за освітньо-професійною програмою «Комп'ютерні системи та мережі»**

**спеціальності 123 «Комп'ютерна інженерія»**

**на тему: «Хмарна система супутникового моніторингу посівів»**

Виконав (-ла):

студент (-ка) IV курсу, групи ІО-64

Пархомчук Олександр Михайлович \_\_\_\_\_

Керівник:

Асистент кафедри ОТ.

Каплунов Артем Володимирович \_\_\_\_\_

Консультант з нормоконтролю:

Професор кафедри ОТ, д.т.н.,

Сімоненко Валерій Павлович \_\_\_\_\_

Рецензент: \_\_\_\_\_

Засвідчую, що у цьому дипломному  
проєкті немає запозичень з праць інших  
авторів без відповідних посилань.

Студент (-ка) \_\_\_\_\_

Київ – 2020 року

**Національний технічний університет України**  
**«Київський політехнічний інститут імені Ігоря Сікорського»**  
**Факультет інформатики та обчислювальної техніки**  
**Обчислювальної техніки**

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 123 «Комп’ютерна інженерія»

Освітньо-професійна програма Комп’ютерні системи та мережі

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_ Сергій Теленик

«\_\_\_» \_\_\_\_\_ 20\_\_ р.

**ЗАВДАННЯ**

**на дипломний проєкт студенту**

**Пархомчук Олександр Михайлович**

1. Тема проєкту «Хмарна система супутникового моніторингу посівів», керівник проєкту Каплунов Артем Володимирович, асистент кафедри ОТ., затверджені наказом по університету від «\_\_\_» \_\_\_\_\_ 20\_\_ р. № \_\_\_\_\_

2. Термін подання студентом проєкту \_\_\_\_\_

3. Вихідні дані до проєкту

4. Зміст пояснювальної записки

5. Перелік графічного матеріалу (із зазначенням обов’язкових креслеників, плакатів, презентацій тощо)

6. Консультанти розділів проєкту\*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
нормоконтроль	Сімоненко В.П., професор, д.т.н.		

7. Дата видачі завдання 01.09.2019 \_\_\_\_\_

\_\_\_\_\_

### Календарний план

№ з/п	Назва етапів виконання дипломного проєкту	Термін виконання етапів проєкту	Примітка
1	Затвердження теми роботи		
2	Вивчення та аналіз завдання	02.12.19-30.12.2019	
3	Розробка алгоритмів	01.01.2020-01.02.2020	
4	Розробка модулю попередньої обробки даних	01.02.2020-04.03.2020	
5	Програмна реалізація системи	05.03.2020-04.04.2020	
6	Оформлення пояснювальної записки	05.04.2020-22.04.2020	
7	Захист програмного продукту	17.05.2020	
8	Передзахист	05.06.2020	
9	Захист	17.06.2020	

Студент

Олександр ПАРХОМЧУК

Керівник

Артем КАПЛУНОВ

### **Анотація**

У дипломній роботі було реалізовано модуль попередньої обробки хмарної системи моніторингу сільськогосподарських посівів на основі безкоштовного програмного забезпечення open data cube .

Розроблені алгоритми дозволяють розширити маску хмар, створювати хмарні та безхмарні композити, які є основою модуля попередньої обробки.

Використовуючи дані супутника sentinel-2 та open data cube, створено систему за допомогою якої можливе отримання та обробка високоякісних даних. Система базується на основі компонентів хмарної платформи Amazon Web Services.

### **Аннотация**

В данной бакалаврской дипломной работе реализовано модуль предварительной обработки облачной системы мониторинга сельскохозяйственных посевов на основе бесплатного программного обеспечения open data cube.

Разработанные алгоритмы позволяют расширить маску облаков, создавать облачные и безоблачные композиты, которые являются основой модуля предварительной обработки.

Используя данные спутника sentinel-2 и open data cube, создана система с помощью которой возможно получение и обработка высококачественных данных. Система базируется на основе компонентов облачной платформы Amazon Web Services.

### **Annotation**

In this work for a Bachelor's Degree, the module of preliminary processing of the cloud system of monitoring of agricultural crops on the basis of free software open data cube was realized.

The developed algorithms allow to expand the mask of clouds, to create cloud and cloudless composites, which are the basis of the pre-processing module.

Using the sentinel-2 satellite data and the open data cube, a system has been created with which high-quality data can be obtained and processed. The system is based on components of the Amazon Web Services cloud platform.

## ВІДОМІСТЬ ДИПЛОМНОГО ПРОЄКТУ

№ з/п	Формат	Позначення	Найменування	Кількість листів	Примітка
1	A4		Завдання на дипломний проєкт	2	
2	A4	ІАЛЦ.467200.000 ВП	Відомість проєкту	1	
3	A4	ІАЛЦ.467200.001 ТЗ	Технічне завдання	3	
4	A4	ІАЛЦ.467200.002 ПЗ	Пояснювальна записка	61	
5	A4	ІАЛЦ.467200.003 Д1	Функціональна схема модуля попередньої обробки	1	
6	A4	ІАЛЦ.467200.004 Д2	Структурна схема	1	
7	A4	ІАЛЦ.467200.005 Д3	Діаграма класів	1	
8	A4	ІАЛЦ.467200.006 Д4	Принципова схема	1	

				ІАЛЦ.467200.000 ВП		
	ПІБ	Підп.	Дата			
Розробн.	Пархомчук О.			Відомість дипломного проєкту	Лист	Листів
Керівн.	Каплунов А.В.				1	1
Консульт.					КПІ ім. Ігоря Сікорського ФІОТ Гр. ІО-64	
Н/контр.	Симоненко В.П.					
Зав.каф.	Стіренко С.Г.					

# **ТЕХНІЧНЕ ЗАВДАННЯ**

до дипломної роботи  
освітньо-кваліфікаційного рівня «бакалавр»

на тему: “Хмарна система супутникового моніторингу  
посіви”

Київ-2020 року

## ЗМІСТ

1.НАЙМЕНУВАННЯ ТА ОБЛАСТЬ ЗАСТОСУВАННЯ.....	2
2.ПІДСТАВИ ДЛЯ РОЗРОБКИ.....	2
3.ЦІЛЬ ТА ПРИЗНАЧЕННЯ ДЛЯ РОЗРОБКИ.....	2
4.ДЖЕРЕЛАМИ ДЛЯ РОЗРОБКИ.....	2
5.ЕТАПИ РОЗРОБКИ.....	3

					ІАЛЦ.467200.001 ТЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		

## 1. НАЙМЕНУВАННЯ ТА ОБЛАСТЬ ПРИЗНАЧЕННЯ

За основу розробки системи взято програмне забезпечення opensource Open Data Cube, яке зараз активно розробляється спільнотами і активно використовується різними урядами. В ODC дуже добре працювати з різними супутниковими знімками які є готовими для аналізу.

## 2. ПІДСТАВИ ДЛЯ РОЗРОБКИ

Підставою для розробки є те, що на сьогоднішній день ще не реалізовано процедури класифікації супутникових даних для отримання карт сільськогосподарських культур на open data cube для України.

## 3. ЦІЛЬ ТА ПРИЗНАЧЕННЯ ДЛЯ РОЗРОБКИ

Цілю даної роботи є розробка модулю попередньої обробки супутникових даних для Open Data Cube, для створення карти моніторингу сільськогосподарських культур використовуючи геодані супутника Sentinel-2.

## 4. ДЖЕРЕЛАМИ ДЛЯ РОЗРОБКИ

Джерелами для розробки є документації по open data cube, python та бібліотекам які використовуються для аналізу супутникових даних, та публікації в Інтернеті по даним питанням.

					ІАЛЦ.467200.001 ТЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		



## 5. ЕТАПИ РОЗРОБКИ

		Дата
1	Ознайомлення з навчальною літературою для програмного забезпечен	07.10.19
2	Отримання навичок з роботою над open data cube	5.05.19
3	Збір та аналіз даних	10.02.20
4	Створення та використання моделі для класифікації	04.04.20
5	Оформлення дипломної роботи	25.05.20

					ІАЛЦ.467200.001 ТЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		

**Пояснювальна записка**  
**до дипломного проєкту**  
**на тему: «Хмарна система супутникового моніторингу**  
**посівів»**

Київ – 2020 року

## ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ.....	7
ВСТУП.....	9
1. ОГЛЯД МЕТОДІВ.....	11
1.1. Crop Watch.....	11
1.2. Crop Monitor.....	14
1.3. Monitoring Agricultural ResourceS (MARS).....	16
1.4. Open Data Cube.....	17
1.5. Sentinel-2.....	20
1.6. Індекс вегетації (NDVI).....	23
1.7. Класифікація методами машинного навчання.....	24
ВИСНОВКИ ДО РОЗДІЛУ 1.....	27
2. РОЗРОБКА АЛГОРИТМІВ.....	28
2.1. Супутникові дані.....	28
2.2. Композитування.....	28
2.3. Безхмарні композити.....	29
2.4. Уточнення маски.....	32
ВИСНОВКИ ДО РОЗДІЛУ 2.....	34
3. РОЗРОБКА СИСТЕМИ.....	35
3.1. Архітектура системи.....	35

					ІАЛЦ.467100.003 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		

3.2. Використання даних.....	36
3.2.1. Векторні дані.....	36
3.3. Розробка функціоналу.....	37
3.3.1. Розробка UML діаграми.....	38
3.4 Функціонал попередньої обробки даних.....	39
3.4.1. Обробка всього датасету.....	42
3.4. Візуалізація геоданих.....	43
3.5. Завантаження даних.....	47
ВИСНОВКИ ДО РОЗДІЛУ 3.....	49
4. ЕКСПЕРЕМЕНТИ ТА ТЕСТУВАННЯ.....	50
4.1. Швидкість роботи.....	50
4.2. Уточнення маски хмари.....	51
4.3. NDVI.....	52
4.4. Класифікація.....	54
ВИСНОВКИ ДО РОЗДІЛУ 4.....	57
ВИСНОВКИ.....	58
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	59
ДОДАТКИ.....	62

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

QGIS – quantum geographic information system .

OSGeo – The Open Source Geospatial Foundation .

ODC - Open Data Cube.

Sentinel-2 – супутник запущений Європейським космічним агентством  
Opensource.

API (application programming interface) – програмний інтерфейс додатку.

NIR (near-infrared spectroscopy) – близько інфрачервона спектроскопія.

SWIR (short-wave infrared) – короткохвильовий інфрачервоний.

Cirrus – димка.

RGB (red, green, blue) – кольорове зображення.

HPC (high performance computing) – високопродуктивні обчислення.

Cloud – хмара.

RF – Random Forest.

DN – digital number.

MLP – multi-layer perceptron.

Copernicus – космічна програма запущена ЄС.

NDVI – нормалізований вегетаційний індекс.

LST - Land surface temperature.

Albedo – фізична величина, що описує здатність поверхні відбивати та розсіювати випромінювання.

FPAR – фракція поглиненого фотосинтетичного випромінювання  
µm - мікрометр

MARS (Monitoring Agricultural ResourceS ) - проект запущений ЄС.

GDAL - Geospatial data abstraction library

QGIS – Quantum geographic information system

ESA – European space agency

NASA - The National Aeronautics and Space Administration

GeoTIFF – вільний стандарт метаданих

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		7

Cloudiness percentage – відсоток хмарності

Cirrus percentage – відсоток димки

Nan percentage – відсоток який немає даних

Метадані – дані які мають певні характеристики

Композит – набір скомбінованих каналів

					ІАЛЦ.467200.002 ПЗ	Арк.
						8
Зм.	Арк.	№ докум.	Підпис	Дата		

## ВСТУП

З кожним роком населення світу збільшується. Зі збільшенням населення в світі людство потребує більше їжі. Проте продовжуючи вирубувати ліси, ми ризикуємо втратити безліч видів тварин, а саме головне втратити зелені легені нашої планети, які забезпечують нас киснем, яким ми дихаємо, але щоб уникнути цієї проблеми, ми можемо навчитися працювати з нашими наявними ресурсами та ретельніше керувати ними і саме в цьому нам можуть допомогти супутники. Різноманітні засоби за допомогою супутників зображують земну кулю точно до розміру одиниці, незалежно від інфраструктури, кордонів, або доброї волі урядів, вони регулярно проводять спостереження, надаючи зображення день у день протягом великого тривалого періоду часу.

Дані насамперед мають бути як повсякденна річ, яка має інформувати та повідомляти, як краще поступити в майбутньому, починаючи від засобів існування та закінчуючи навколишнім середовищем. Проблема полягає в тому як нам правильно навчитись використовувати ці дані, адже на даний момент кожен користувач має доступ до супутникових даних. Якби кожна країна використовувала супутникові дані з метою покращення, то можна було б вирішити безліч проблем.

На сьогоднішній день дистанційне зондування землі ефективно використовується у всьому світі, адже з кожним роком кількість супутників які обертаються навколо землі стрімко зростає. Тому виникає необхідність в обробці цих геоданих, адже більшість даних знаходяться у відкритому доступі.

Дистанційне зондування землі сільськогосподарських культур має проводитись на основі супутникових даних високої роздільної здатності, адже це є дуже важливим компонентом для отримання більш чітких результатів.

Саме за допомогою sentinel-2 ми можемо побачити як вирубка лісів впливає на середовище, також ми можемо відрізнити здорове дерево від мертвого завдяки його інфрачервоним слідам, між пишним лугом і сухим, або між багатим і низьким урожаєм, що робить різницю між дією та голодом

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		9

мільйонів людей. Завдяки дистанційному зондуванню землі ми можемо бачити незаконні вирубки лісів у найвіддаленіших куточках світу.

Завдяки безкоштовному доступу до даних за допомогою сервісу Copernicus та високої просторової роздільної здатності супутник sentinel-2 допомагає нам слідкувати за цим, тому завдяки моніторингу ми можемо розробити моделі які будуть здатні спрогнозувати як поводитися з природними ресурсами, такими як вода, гній, та ґрунт, тобто де вони потрібні, а де ні, також ми можемо спрогнозувати де потрібно робити зрошення, де уникати надмірних добрив наших полів, або коли і де потрібно засіювати поля.

За основу було взято програмне забезпечення яке на даний момент є у відкритому доступі і зараз активно розробляється спільнотами і використовується різними урядами.

На даний момент на open data cube ще не реалізовано процедури моніторингу сільськогосподарських посівів за допомогою супутникових даних sentinel-2 для отримання сільськогосподарських культур. Тому метою роботи є реалізувати нормальну вичитку даних з open data cube у часові ряди для створення подальшої системи супутникового моніторингу посівів.

Використовуючи систему на хмарній платформі, ми можемо подолати проблеми із завантаженням даних, зберіганням, та обробкою великих даних. За допомогою Amazon Web Services (AWS) таких як S3 та EC2 ми можемо мати легкий та швидкий доступ до даних та обчислювальних ресурсів.

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		10



# РОЗДІЛ 1

## ОГЛЯД МЕТОДІВ

### 1.1 Crop watch

В даному розділі розглядається система моніторингу сільськогосподарських посівів в Китаї. Компанія Cropwatch створена в 1998 році. Даний сервіс використовує геодані дистанційного зондування землі високого (30метрів та вище) та низького (250-1000метрів) розширення для різних типів обчислення та подання різних показників моніторингу сільськогосподарських посівів. Моніторинг виробництва продуктів харчування та раннє попередження передбачені лише для Китаю. Незважаючи на це компанія crop watch також проводить оцінку посівів сільськогосподарських культур, оцінку врожаю та виробництво врожаю у світовому масштабі, але оцінка площі врожаю проводиться лише для 6 країн (США, Аргентина, Бразилія, Франція, Німеччина та Україна). На рисунку 1.1 ми можемо бачити як архітектура працює.

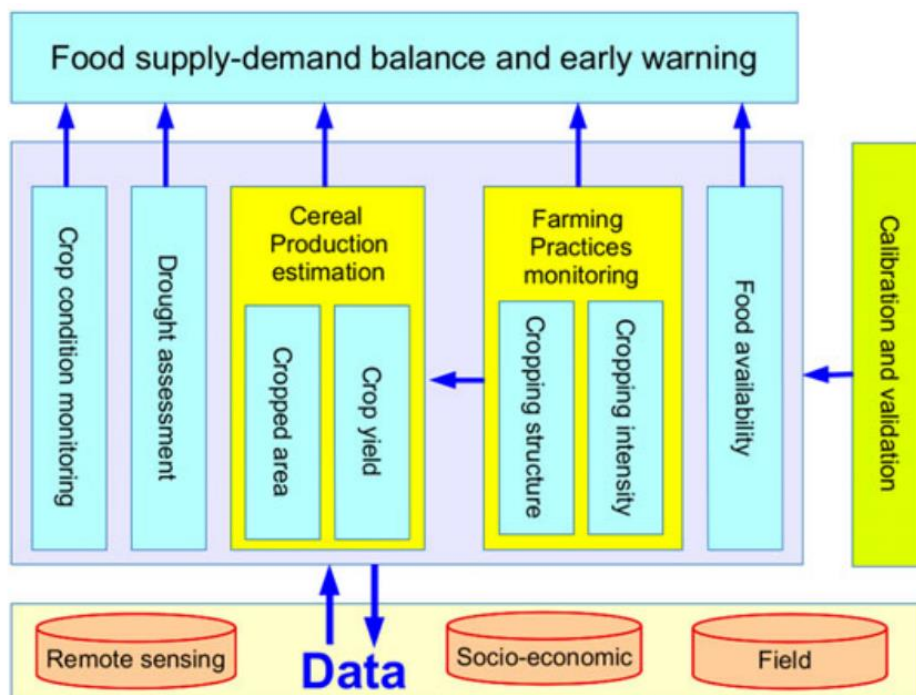


Рис. 1.1 CropWatch архітектура системи[1]

Система використовує геодані з різних джерел дистанційного зондування, у разі потреби дані супутників завантажуються та попередньо обробляються за допомогою створеними методами та зберігаються в базі даних. Для покращення результатів компанія використовує власні методи обробки даних, один з методів є усунення шуму. Принцип підготовки геоданих та дистанційного зондування землі компанії cropwatch більш детально описує таблиця 1.1.

Таблиця 1.1 CropWatch ДЗЗ та підготовка даних[1]

Територія застосування	Супутники	Роздільна здатність	Препідготовка даних	Продукти
Китай	FY-3a MERSI	250/1000м	Геометрична корекція, калібрування DN-зображень на TOA коефіцієнт відбиття, корекція атмосфери, ідентифікація хмари та маскування	NDVI, LST, каналні віддзеркалення
	TERRA/AQUA MODIS	250/1000 м	Географічна корекція; калібрування радіацією; зміна розміру зображення; мозаїка, хмарна ідентифікація; та атмосферна корекція	NDVI, albedo, LST
	ENVISat ASAR	30м	Геометрична корекція, калібрування сйва та спектр скорочення	Коефіцієнт зворотнього розсіювання
	Radarsat-1	30м	Геометрична корекція, калібрування сйва та спектр скорочення	Коефіцієнт зворотнього розсіювання , текстури

Таблиця 1.1(продовження) CropWatch ДЗЗ та підготовка даних[1]

Територія застосування	Супутники	Роздільна здатність	Препідготовка даних	Продукти
Глобальне	Landsat TM/ETM	30м	Геометрична корекція, калібрування DN-зображень на TOA коефіцієнт відбиття, корекція атмосфери, ідентифікація хмари та маскування	коефіцієнт відбиття смуг
	HJ-1 CCD	30м	Геометрична корекція, калібрування DN-зображень на TOA коефіцієнт відбиття, корекція атмосфери, ідентифікація хмари та маскування	коефіцієнт відбиття смуг
	IRS P6 AWIFS	56м	Геометрична корекція, калібрування DN-зображень на TOA коефіцієнт відбиття, корекція атмосфери, ідентифікація хмари та маскування	коефіцієнт відбиття смуг
	TERRA/AQUA MODIS	1000м	мозаїку, перепроєктування	NDVI, FPAR, LST
	NOAA AVHRR	1000м	Хмарна ідентифікація та маскування; калібрування; атмосферна корекція; виведення LST з теплових смуг; і геометричні виправлення	NDVI, NDVI, LST

Також дана компанія пропонує свій ієрархічний підхід для моніторингу сільськогосподарських посівів який передбачає різні екологічні показники на різних масштабах, використовуючи об'єднану інформацію для оцінки регіональних та національних умов. Дана ієрархія зображена у вигляді піраміди на рисунку 1.2.

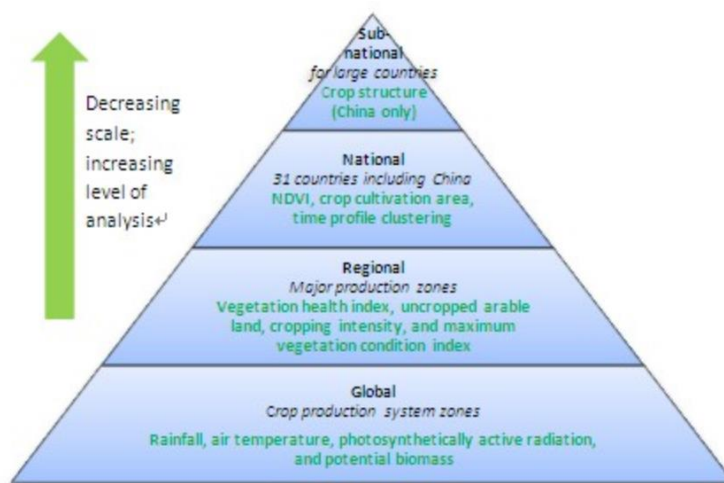


Рис. 1.2. Ієрархічний підхід CropWatch[2]

Для того, щоб аналізувати ці дані компанія визначає кожен рівень конкретними показниками, пристосованими до його масштабу, тоді як наступний більш широкий рівень забезпечує загальну основу. Використовуючи ці результати на чотирьох рівнях.

## 1.2 Crop Monitor

Crop Monitor – це проект від NASA який створений для об’єднання міжнародних партнерів. Даний проект збирає інформацію про всі чотири види сільськогосподарських культур таких, як рис, пшениця, кукурудза та соя. Система базується на поєднанні національних та регіональних даних порівнюючи із даними які надаються дистанційним зондуванням землі, спираючись на експертні знання та досвід міжнародних експертів у галузі сільського господарства з усього світу. На рисунку 1.3 ми можемо бачити карту сільськогосподарського моніторингу посівів від Crop Monitor.

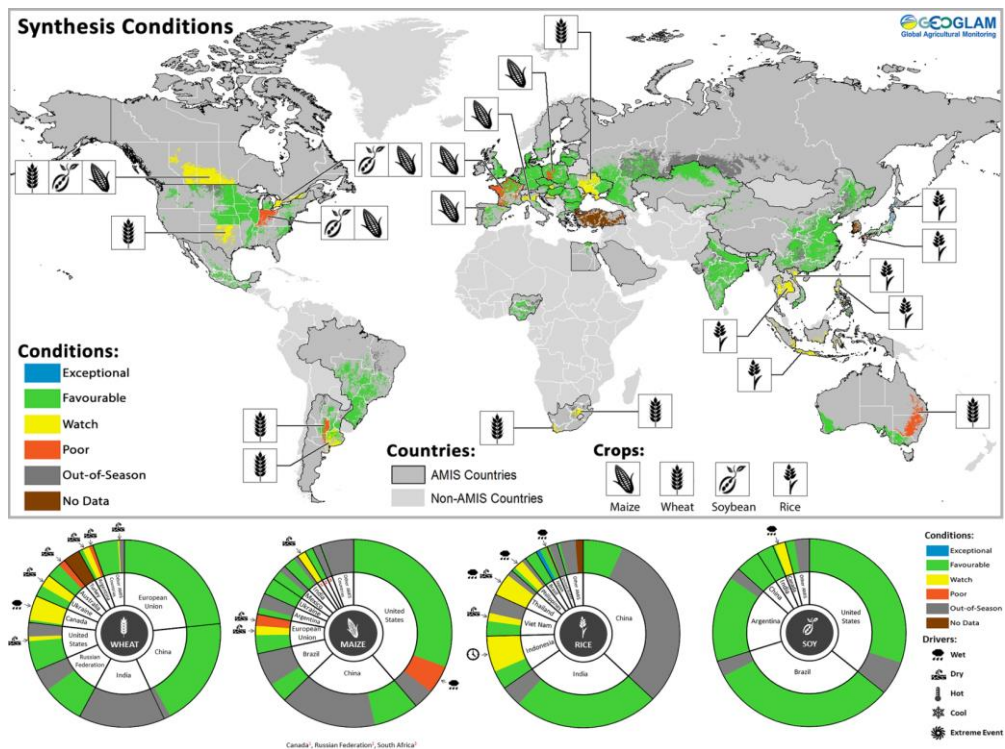


Рис. 1.3. Карта станів сільськогосподарських культур, які входять до інформаційної системи сільськогосподарського ринку[3]

Дана карта охоплює країни які входять в велику двадцятку G20, тобто країнами з найбільшою економікою в світі, також карта містить сім основних країн-виробників та експортерів.

У цьому проекті поєднуються глобальні індекси які засновані на дистанційному зондуванні землі, з масками для врожаю та календарями для врожаю, для вирішення проблем з врожайністю, та знаходження проблемних регіонів які можуть впливати на подальший вегетаційний період. На рисунку 1.4 можна побачити як працює система сільськогосподарського моніторингу[3].

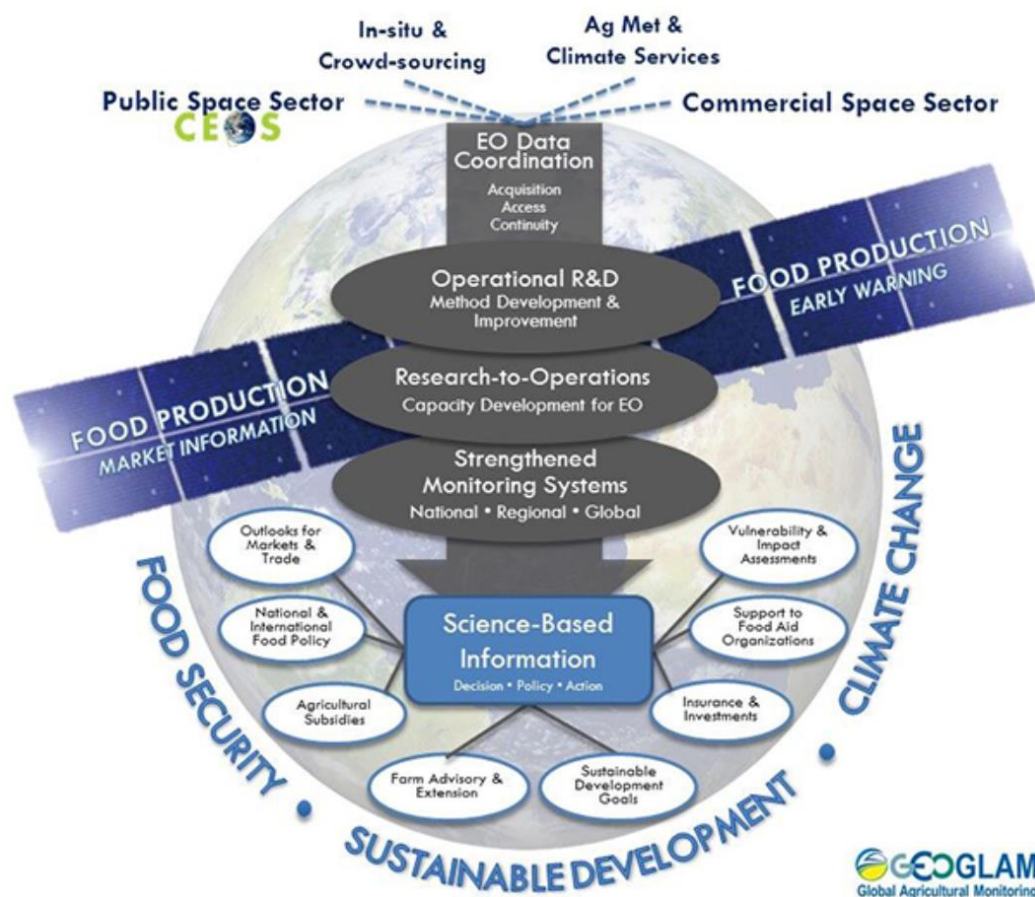


Рис 1.4 Система глобального сільськогосподарського моніторингу Crop Monitor[4]

Даний проект використовує супутникові дані: MODIS, ATOVS, SMOS, SMAP, ASCAT, TIRS, TRMM.

### 1.3 Monitoring Agricultural ResourceS (MARS)

Monitoring Agricultural ResourceS (MARS) – це європейський проект за допомогою якого здійснюється сільськогосподарський моніторинг посівів. Цей проект розробляє різні методи, інструменти для використання в рамках моніторингу сільського господарства, що застосовується для Європи, Африки та інших регіонів світу[5]. Моніторинг врожаю проводиться для надання щомісячних результатів, що прогнозують врожайність сільськогосподарських культур та за допомогою комп'ютерних моделей створюються моделі подальшого росту та розвитку культур. Отримані дані з моделей передають в науково-дослідницький центр, де всі ці дані аналізуються та створюється

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		16

статистика за допомогою системи прогнозування, яка допомагає в подальшому визначити які сільськогосподарські культури зможуть вижити в даному сезоні, які регіони зможуть постраждати через зменшення кількості опадів. Це і є основним принципом роботи проекту MARS. Також в рамках сільськогосподарської політики ЄС в проекті постійно вдосконалюються методи та настанови для вимірювання площ посівів, визначення типу сільськогосподарських культур, особливостей ландшафту та оцінки впливу на навколишнє середовище.

Моніторинг сільськогосподарських посівів здійснюється за допомогою дистанційного зондування землі[6].

## 1.4 Open Data Cube

Моніторинг сільськогосподарських посівів це задача з великим об'ємом даних, тому щоб реалізувати цю задачу її потрібно виконувати в хмарі. Для реалізації цього було використано таке програмне забезпечення, як Open Data Cube.

Open Data Cube (ODC) – це набір безкоштовного програмного забезпечення для аналізу геопросторових даних та спостереженням за земним покривом, багатьох супутників та інших систем збору даних. ODC значно полегшує обробку даних та дозволяє зберігати великі об'єми даних. Завдяки ODC ми можемо вибрати територію яка нас цікавить і подивитись як вона змінювалась за певний період часу[7].

Основна характеристика ODC полягає в тому, що завдяки цій платформі можна робити унікальні спостереження, які відрізняються від інших методів геоборобки даних. Основною перевагою ODC є:

- Колекція великих обсягів геопросторових даних для спостереження за земним покривом
- API на основі Python для високопродуктивних запитів доступу даних
- Надає вченим та іншим користувачам легку можливість виконувати дослідницький аналіз даних

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		17

- Зберігати масштабовану обробку континенту
- Відстежувати всі дані, що містяться, для забезпечення контролю якості та оновлення

Ядро ODC знаходиться між супутниковими даними та користувачем. Використовуючи хмарну систему Open Data Cube ми значно полегшуємо роботу при обробці даних, тобто нам не потрібно викачувати знімки для того, щоб їх обробляти[8]. Реалізація Open Data Cube складається з даних, інфраструктури, та додатків.

Використання даних базується на файлах або каталогах у нашому випадку весь датасет зберігається в форматі NetCDF. Для того щоб працювати з цим форматом використовуються бібліотеки такі як GDAL та rasterio, всі дані зазвичай зберігаються на серверах Amazon S3.

Для зберігання даних Open Data Cube використовує базу даних PostgreSQL, для зберігання готових для аналізу супутникових даних. Інфраструктура побудована таким чином, що користувач може зчитувати дані та місцеположення, не вимагаючи конкретної вказівки знаходження файлу.

Програмне забезпечення на основі Open Data Cube є бібліотека Python. За допомогою цієї бібліотеки користувач може легко керувати даними, приймати дані, оптимізувати для більшої продуктивності системи, запитувати дані, та використовувати багато інших функцій, які полегшують роботу з даними.



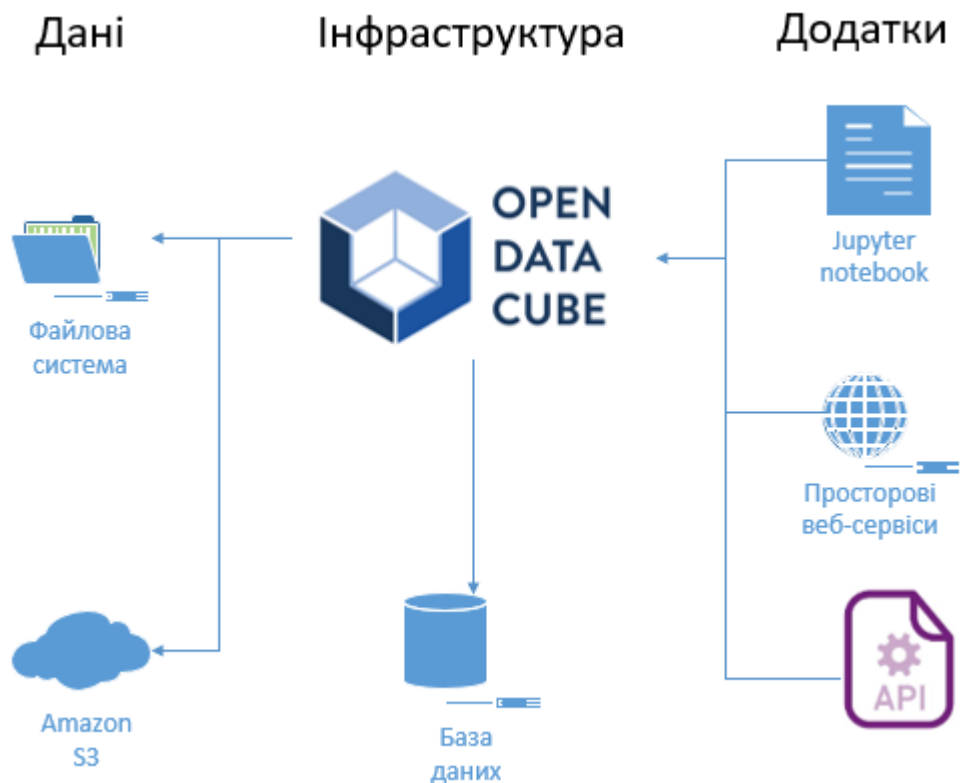


Рис. 1.5 Реалізація системи на технічному рівні[9]

Використовуючи open data cube ми можемо вибрати потрібні геодані які нас цікавлять, та аналізувати їх. Головний принцип open data cube це полегшити управління з великими об'ємами даних, не вимагаючи збереження даних, та максимізувати обробку геоданих які ми отримали, для подальших рішень. Таким образом open data cube робить можливою роботу, яку неможливо було представити кілька років тому.

Завдяки сервісам Amazon дуже зручно користуватися open data cube, адже всі дані знаходять на серверах, що значно прискорює роботу. На рисунку 1.6 ми можемо бачити карту світу де Open Data Cube є запущений та країни які знаходяться на різних етапах впровадження.



Рис. 1.6 Використання ODC у світі[10]

Оскільки з кожним роком нові покоління дистанційного зондування землі створюють більші об'єми даних, та для обробки цих даних потребується більше ресурсів, тому архітектурне рішення системи допомагає значно знизити потребу в більшій кількості обчислювальних даних, завдяки оптимізації системи. Також ми можемо бачити що кількість країн які зацікавлені у використанні Open Data Cube є досить великою.

## 1.5 Sentinel-2

В даній роботі для класифікації використовується датасет Sentinel-2. Sentinel-2 – це Європейська широкомасштабна місія запущена у рамках програми «Коперник» з багатозоровим спектром зображень, для моніторингу земної поверхні. Дана місія відбулася завдяки тісній співпраці між ESA[11], Європейською Комісією для того щоб продемонструвати технологічну досконалість Європи. У розробці цієї місії було залучено понад 60 компаній Перший супутник sentinel-2A був запущений 23 червня 2015року, другий супутник sentinel-2B був запущений 7 березня 2017року. Ці супутники рухаються по протилежні боки орбіти та рухаються поетапно при 180° один до одного[12][13]. Характеристики місії sentinel-2 наведені в таблиці 5.1.

Табл. 1.2 Характеристики місії sentinel-2[14]

Sentinel-2A запуск	23 червня 2015 року, Вега з Куру, Французька Гвіана
Sentinel-2B запуск	Березень 2017 року, Вега з Куру, Французька Гвіана
Орбіта	Сонцесинхронний на висоті 786 км, середній місцевий сонячний час у низхідному вузлі: 10:30 (оптимальне освітлення Сонця для отримання зображення)
Геометричний час перегляду	П'ять днів від дво супутникового сузір'я (на екваторі)
Термін експлуатації	Сім років (переносить витратні матеріали протягом 12 років: 123 кг пального, включаючи закінчення терміну експлуатації)
MSI (багатоспектральний канал)	MSI, що охоплює 13 спектральних смуг (443–2190 нм), шириною ширини 290 км та просторовим дозволом 10 м (чотири видимих та близько інфрачервоних смуг), 20 м (шість червоних крайових та короткохвильових інфрачервоних смуг) та 60 м (три смуги корекції атмосфери).
Приймальні станції	Дані MSI: передаються через X-діапазон до основних наземних станцій Sentinel та по лазерному каналу через EDRS. Дані телекомунікацій та телеметрії: передані з та в Кіруну, Швеція
Основні програми	Сільське господарство, ліси, зміна землекористування, зміна земельного покриву. Картографування біофізичних змінних, таких як вміст хлорофілу в листі, вміст у листі води, індекс площі листя; моніторинг прибережних та внутрішніх вод; картографування ризиків та катастроф
Місія	Керуються, розробляються, експлуатуються та експлуатуються різними установами ESA
Фінансування	Держави-члени ЄКА та Європейський Союз
Основні підрядники	Airbus Defence & Space Germany for the satellite, Airbus Defence & Space France for the instrument
Співпраця	CNES: Оптимізація якості зображення під час введення в орбіту DLR: Optical Communication Payload (provided in kind) NASA: cross calibrations with Landsat-8

На борту супутників sentinel-2 знаходиться мультиспектральні інструменти (MSI) за допомогою якого проводяться виміри відбитого світла землі у 13 видимих, близькому інфрачервоному(VNIR) і інфрачервоному з короткими хвилями (SWIR) спектральних діапазонах. За допомогою цих

інструментів супутник збирає дані та обробляє їх. Можливості мультиспектральної візуалізації роблять місію sentinel-2 настільки надзвичайною, що користувачі можуть бачити в зміні росту рослинності, з безпрецедентною точністю[15]. Мультиспектральні канали sentinel-2 наведені в таблиці під номером 5.2.

Табл. 1.3. Мультиспекторні канали sentinel-2[16]

<b>Канали Sentinel-2</b>	<b>Центральна довжина хвилі (μм)</b>	<b>Роздільна здатність(м)</b>	<b>Ширина смуги каналу(нм)</b>
Канал 1 - Прибережні аерозолі	0.443	60	20
Канал 2 – Синій	0.490	10	65
Канал 3 – Зелений	0.560	10	35
Канал 4 – Червоний	0.665	10	30
Канал 5 – Вегетаційний червоний край	0.705	20	15
Канал 6 – Вегетаційний червоний край	0.740	20	15
Канал 7 – Вегетаційний червоний край	0.783	20	20
Канал 8 – NIR	0.842	10	115
Канал 8A – Вузький NIR	0.865	20	20
Канал 9 – Водяна пара	0.945	60	20
Канал 10 – SWIR – Cirrus	1.375	60	20
Канал 11 – SWIR	1.610	20	90
Канал 12 - SWIR	2.190	20	180

За допомогою технології мультиспекторних сенсорів, супутник може передавати дані через лазер до інших супутників, що знаходяться на геостаціонарній орбіті. Використання цієї магістралі космічних даних дозволяє дуже швидко передавати великі обсяги даних, щоб зробити інформацію більш доступнішою для користувачів. Знімок однієї й тієї самої території проводиться раз в 5днів.

За один оберт навколо землі супутник створює полосу даних ширина якої складає 290 кілометрів, ці полоси є ширшими за будь який інший супутник в цьому класі. Обсяг даних однієї полоси складає 1.6 ТБ при просторовому розрізненні 10м. Всі дані супутника sentinel-2 поставляються до наземної служби «Коперник», які потім зберігаються у відкритому доступі і кожен користувач може скористуватися ними. Завдяки цим даним ми можемо бачити різні види рослин та їх хвороби, індекс площі листя, вміст хлорофілу який має важливе значення для точного землеробства. За допомогою хронології часу ми можемо отримати точний моніторинг вегетаційного періоду, отримати підтримку у важливих областях, які приносять користь суспільству таких як продовольча безпека. Також ми можемо моніторинг зміни земного покриву які були пошкодженні, наприклад, пожарами[17].

## 1.6 Індекс вегетації (NDVI)

За допомогою формули індексу вегетації ми можемо слідкувати за життям рослин. Індекс вегетації – це відношення між ближнім інфрачервоним спектром (NIR) та червоним спектром, поділене на їх суму. Вегетаційний індекс обчислюється за формулою (1.1):

$$NDVI = \frac{(NIR-RED)}{(NIR+RED)}, \quad (1.1)$$

Де NIR – це інфрачервоний спектр.

RED – червоний спектр.

Принцип роботи вегетаційного індексу полягає в тому, щоб спостерігати скільки червоного світла поглинається та відбивається. За допомогою цієї формули ми можемо відрізнати здорові рослини від хворих. Здорові рослини поглинають більшу частину хвиль спектру, а коли рослини є хворими, спектри світла будуть відбиватися. Отримані результати для декількох пікселів можуть

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		23

бути використані для подальшої візуалізації, чи аналізу місцевості за зміною фотоситетичної рослинності на територіях.

## 1.7 Класифікація методами машинного навчання

В даному розділі розділі розглядаються методи машинного навчання для земного покриття.

Машинне навчання (Machine Learning) – це підрозділ штучного інтелекту, основою якого є обробка великої кількості прикладів, виявлення закономірностей для того щоб спрогнозувати дані. Існують два типи навчання, це навчання з вчителем та навчання без вчителя.

В даній роботі для класифікації було використано навчання з вчителем, адже це один з найосновніших видів машинного навчання. Цей алгоритм навчається на навчальній вибірці, та згодом за допомогою цієї вибірки виконується подальша обробка даних.

Використовуючи безкоштовне програмне забезпечення (Quantum GIS) було створено вибірку для подальшої класифікації сільськогосподарських посівів. В цій вибірці кожен елемент мав ідентифікатор, клас та назву.

Класифікація сільськогосподарських посівів проводиться за допомогою двох технологій машинного навчання: випадковий ліс (random forest) та багат шаровим перцептроном (multilayer perceptron). Кожен метод має свої переваги та недоліки. Також точність класифікації залежить не тільки від методу класифікації, а й від якості навчальних даних, та звичайно параметрів класифікатора.

Випадковий ліс (RF - Random forest) – це метод машинного навчання для класифікації. Він є одним з найкращих алгоритмів для прогнозування, бо при цьому він абсолютно нічого не пояснює, як побудовані дані. Цей метод створює ліс з декількома деревами рішень - чим більше створено дерев, тим надійнішим буде прогнозування, та в результаті чого буде більша точність [18].

Коли навчальний набір для поточного дерева проводиться шляхом відбору проб із заміною, приблизно одна третина випадків залишається поза

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		24

вибіркою. Ці дані використовуються для отримання неупередженої оцінки помилки класифікації, оскільки дерева додаються до лісу.

Після побудови кожного дерева всі дані запускаються по дереву, та обчислюється близькість для кожної пари витків. Якщо два витки займають один і той же кінцевий вузол, їх близькість збільшується на одиницю. В кінці алгоритму близькість нормалізується діленням на кількість дерев. Близькість використовується для заміни відсутніх даних.

Переваги випадкового лісу полягають в тому, що класифікатор може використовуватись як для задач класифікації так і для задач регресії. Класифікатор буде обробляти пропущені значення і підтримувати їх точність коли більша частина даних є відсутньою.

Коли ми маємо більше дерев тоді перша випадкова ситуація не буде відповідати моделі, вона може обробляти великі набори даних з великою розмірністю.

Багатошаровий перцептрон (MLP – multi-layer perceptron) – це нейронна мережа, яка має простий алгоритм, який є призначений для бінарної класифікації, тобто він передбачає, належить певна категорії даному класу чи ні. Персептрон є лінійним класифікатором, тобто алгоритм класифікує вхід, розділяючи дві категорії прямою лінією. Це найпростіший представник нейронних мереж, який описує один нейрон[19].

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		25

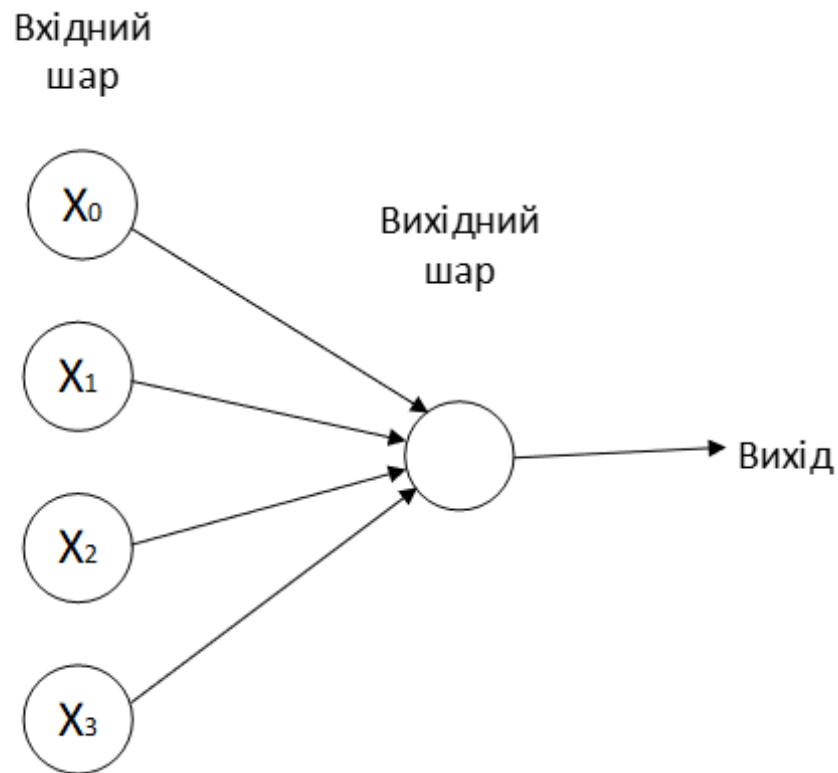


Рис. 1.7 Приклад роботи перцептрону[20]

Перцептрон можна описати формулою(1.2):

$$f(x, w, b) = \begin{cases} 1, & \text{якщо } \sum_{i=1}^m w_i * x_i + b > 0, \\ 0 & \text{інакше} \end{cases} \quad (1.2)$$

Де  $X$  – вхідні вектори.

$W$  – вага вектору.

$b$  – зміщення, або поріг спрацювання, якщо перенести його вправо

Виходячи з формули можна сказати, що перцептрон активується, якщо сума входів помножена на вагу є більшою порогу спрацювання.

Для класифікації використовується бібліотека `scikit-learn`. Саме ця бібліотека є простим та зручним інструментом, за допомогою якої ми з легкістю можемо користуватися різними методами машинного навчання та аналізу даних. Завдяки цій бібліотеці ми з легкістю можемо взаємодіювати з різними модулями. Також за допомогою цієї бібліотеки ми можемо стандартизувати дані та підготувати дані для машинного навчання.

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		26



## ВИСНОВКИ ДО РОЗДІЛУ 1

Проаналізовано існуючі системи агромоніторингу, їх недоліки та запропоновано високорівневу архітектуру розподіленої системи на основі компонентів хмарної платформи Amazon Web Services та програмного забезпечення DataCube. На основі цих даних було вирішено розробити алгоритми які є наближеними до тих які використовуються на даний момент в світі.

					ІАЛЦ.467200.002 ПЗ	Арк.
						27
Зм.	Арк.	№ докум.	Підпис	Дата		

## РОЗДІЛ 2

### РОЗРОБКА АЛГОРИТМІВ

#### 2.1 Супутникові дані

Для розв'язання поставленої задачі використано супутникові знімки sentinel-2. В якості розробки власних алгоритмів які є наближеними до тих, що використовуються на даний момент в світі.

#### 2.2 Композитування

Композитування грає важливу роль у дослідженні та аналізі супутникових даних, оскільки завдяки ньому можна одержати інформацію про конкретний піксель знімку. Найбільш важливим є правильна комбінація каналів зображення, оскільки інформація що надходить до джерела може бути не відсортованою по каналам. Також композитинг знімків дає можливість застосовувати різні методи розпізнавання образів або регресійні методи машинного навчання, де на вхід завжди подається стек каналів і на вихід одержується по одному каналу за один прохід алгоритму[21].

Увесь композитинг каналів, що надавались веб-сервісом scihub copernicus був реалізований за допомогою прикладної бібліотеки numpy, що дозволяє швидко та комфортно керування масивами даних.

Кожний канал мультиспектрального зображення може відображатися по одному каналу в різних відтінках сірого кольору. При комбінації трьох каналів ми можемо отримати кольорове зображення.

Проаналізувавши дані sentinel-2 ми можемо бачити, що моніторинг земного покриву проводиться в 13 мультиспектральних каналах, тобто готових кольорових зображень таких, як RGB, супутник нам не пропонує. Тому, щоб отримати кольорове зображення таке як RGB, ми створюємо композит з трьох каналів, а саме з червоного, зеленого, синього.

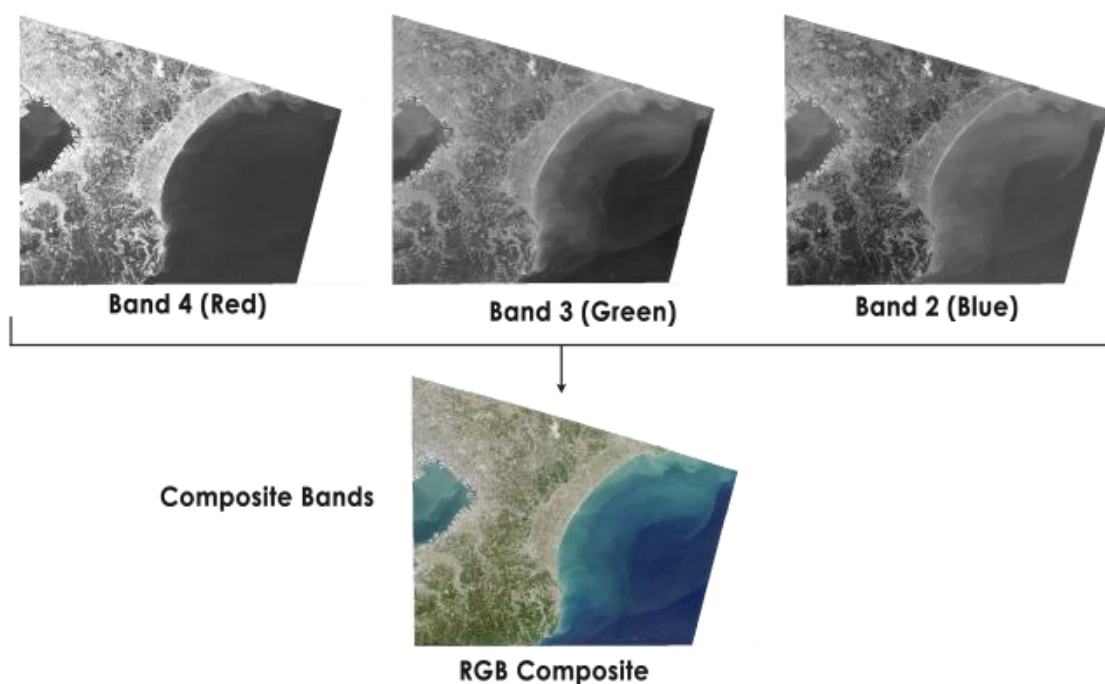


Рис. 2.1. Композитування знімків sentinel-2

Коли ми поєднуємо ці три канали в результаті ми маємо кольорове зображення з кольором кожного пікселя, що визначається комбінацією RGB різної яскравості. Отриманий композит демонструє нам зображення яке людина бачить своїм оком.

Створений композит є готовий для подальших досліджень, кожен знімок маж просторове розрізнення 10м.

### 2.3 Безхмарні композити

В даному розділі описується процес створення безхмарних композитів на основі даних sentinel-2.

Безхмарні композити є одним з попередніх етапів обробки будь яких геоданих.

Отримавши кольорові знімки можна побачити, що головною проблемою аналізу геоданих є захмарені знімки. Нажаль місією sentinel-2 не передбачено прибирання хмар, але з іншого боку sentinel-2 пропонує нам класифікацію земної поверхні, тобто завдяки мультиспекторним каналам створюється класифікація земної поверхні, яка подається нам у вигляді маски.

Нижче приведена таблиця класифікації об'єктів маски, яку пропонує нам sentinel-2.

Табл. 2.1. Класифікація об'єктів маски[22]

Значення	Класифікація
0	Немає даних
1	Дефектні об'єкти
2	Область темних пікселів
3	Тінь хмар
4	Рослинність
5	Не рослинність
6	Вода
7	Не класифіковано
8	Хмари середньої прозорості
9	Не прозорі хмари
10	Димка
11	Сніг

Маска яку нам пропонує sentinel-2 не є ідеальною, але з іншого боку завдяки цій масці ми можемо бачити де знаходяться хмари на супутникових знімках, що значно полегшує роботу.

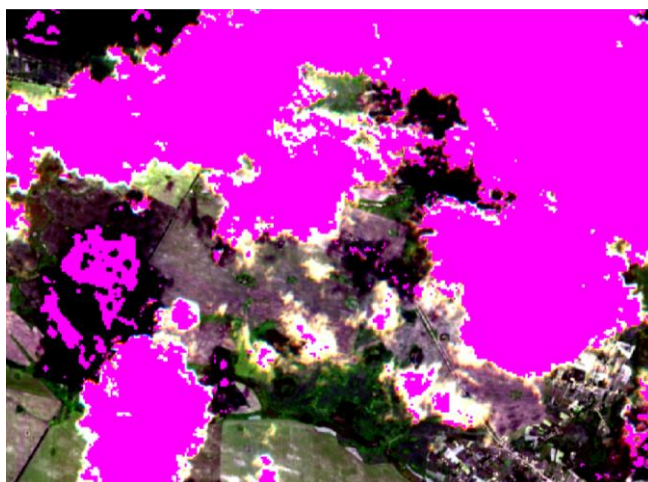


Рис. 2.2 Недооцінка маски хмари

На малюнках ми можемо бачити, що ця маска також не є ідеальною, зазвичай ця маска не завжди охоплює повністю хмару, та й інколи маска покриває територію де немає хмари.

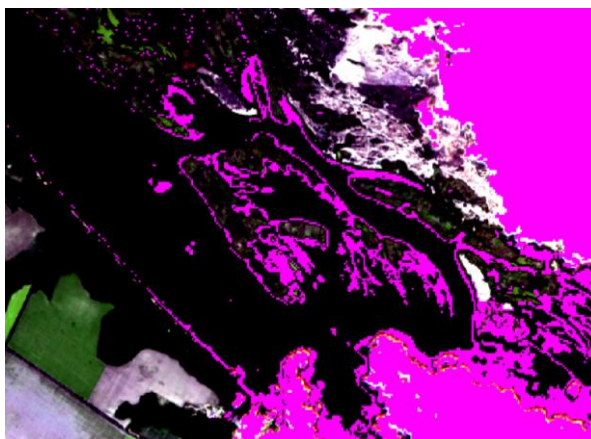


Рис. 2.3 Переоцінка маски хмар

Інколи трапляється таке, що хмари покривають більше 80% знімків і це значно погіршує в кінці результат, так як sentinel-2 робить знімок однієї й тієї самої території кожні 5 днів. Тому це є проблемою для отримання більш точних результатів класифікації.

Проаналізувавши ці дані, було розроблено алгоритм, який дозволяє нам вирізати хмари за допомогою маски хмар та замінити порожні пікселі на найближчі дати. Для розробки цього алгоритму була використана класифікація об'єктів sentinel-2.

Табл. 2.2 Мітки класів для уточнення хмар

Значення	Класифікація
0	Немає даних
1	Дефектні об'єкти
2	Область темних пікселів
3	Тінь хмар
7	Не класифіковано
8	Хмари середньої прозорості
9	Не прозорі хмари
10	Димка
11	Сніг

## 2.4 Уточнення маски

В даному розділі розглядається процес розширення маски sentinel-2. Зазвичай ця маска подається в 60 метрах, тобто 1 піксель дорівнює 60 метрам. Це також ускладнює роботу, адже для більш точних даних ми використовуємо 10 метрів на 1 піксель. Тобто щоб отримати більш чітку маску відповідної зони було використано алгоритм **Dilation**[23].

Маска яку пропонує нам sentinel-2 не є якісною, та вона не охоплює всю хмару і з'являються дефекти на краях.

Проаналізувавши ці результати, було вирішено застосувати розширення маски, а саме алгоритм Dilation.

Алгоритм працює таким чином є дві множини, множина A і множина B як структурний елемент. Даний елемент обходить множину A, тим самим збільшуючи кожен раз на зрушення s. Алгоритм dilation, написаний як

$A \oplus B$ , тобто це є сукупність всіх зрушень, які задовольняють:

$$A \oplus B = \{s | (B)s \cap A\}$$

Приклад роботи dilation:

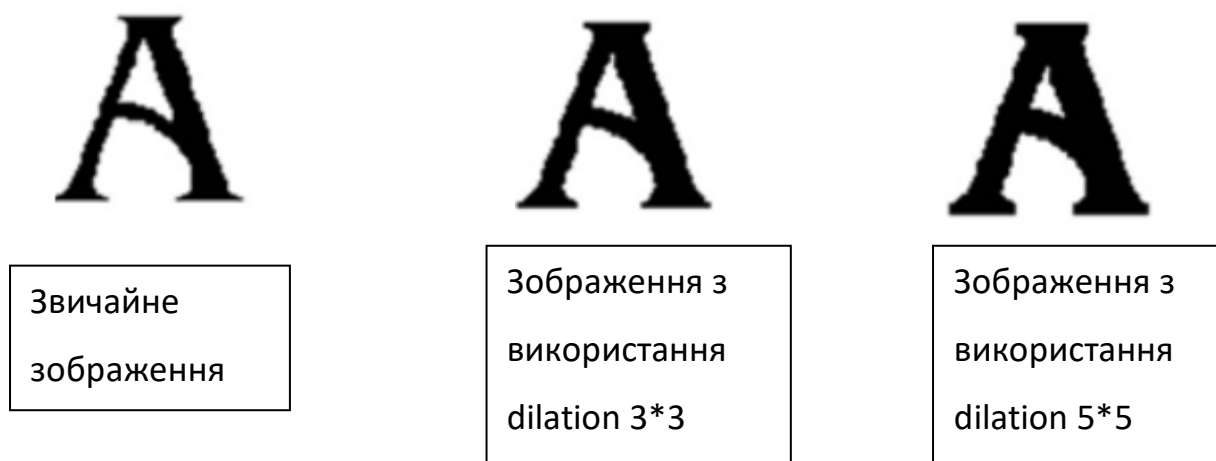


Рис. 2.4 Приклад використання dilation

Безхмарні композити є вже готовими даними для аналізу. На основі цих даних будуть проводитись подальші дослідження. Використовуючи dilation не можна сказати, що за допомогою цього алгоритму можна прибрати абсолютно всі хмари. Прибирання повністю всіх хмар не є дуже легкою роботою, тому з цією проблемою стикаються всі хто працюють з геоданими. Але застосувавши цей алгоритм, ми мінімізуємо кількість даних які ми можемо втратити для більш точного аналізу.

					ІАЛЦ.467200.002 ПЗ	Арк.
						33
Зм.	Арк.	№ докум.	Підпис	Дата		

## ВИСНОВКИ ДО РОЗДІЛУ 2

Дані які ми отримуємо з супутника не є обробленими, тому для подальшої обробки нам потрібно їх підготувати. Тому було вирішено розробити архітектуру модуля попередньої обробки геопросторової інформації.

Розроблені алгоритми дозволяють розширити маску хмар, створювати хмарні та безхмарні композити, які є основою модуля попередньої обробки.

					ІАЛЦ.467200.002 ПЗ	Арк.
						34
Зм.	Арк.	№ докум.	Підпис	Дата		



## РОЗДІЛ 3

### РОЗРОБКА СИСТЕМИ

#### 3.1 Архітектура системи

Для створення хмарної системи моніторингу посівів була розроблена архітектура, яка поділяється на 3 головних модулі: дані, функціональна система та інтерфейс. Система хмарного моніторингу посівів потребує великої кількості обчислювальних ресурсів, тому було вирішено розробити дану систему в хмарі за допомогою сервісів Amazon. Структурна схема системи знаходиться в додатку 1.

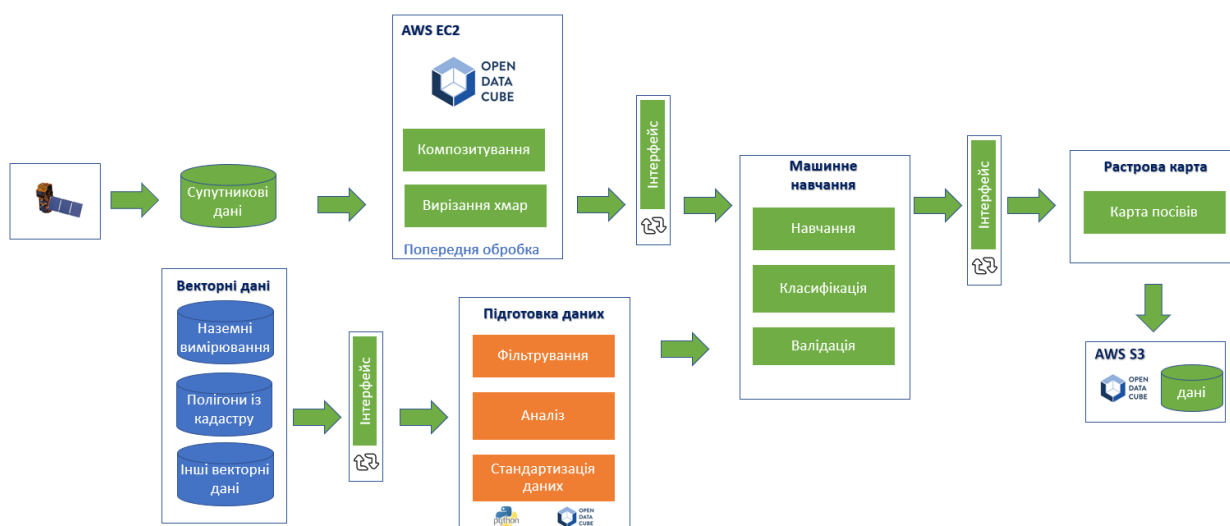


Рис. 3.1. Архітектура хмарної системи супутникового моніторингу посівів[24]

Основним функціоналом обробки даних цієї діаграми є: завантаження та попередня обробка супутникових зображень, тренування, класифікація та валідація за допомогою машинного навчання, та отримання кінцевого результату, тобто карти класифікації посівів[25].

Дана архітектура допомагає значно прискорити роботу супутникового моніторингу посівів. Використовуючи сервери AWS EC2, ми виконуємо всю обробку даних в хмарі. Функціональна робота системи знаходиться в додатку 2. За допомогою цього сервісу ми значно полегшуємо проведення обчислень в

хмарі. Тобто зручний веб-інтерфейс з легкістю дозволяє нам отримати доступ до обчислювальних ресурсів з мінімальними зусиллями, та повним контролем ресурсів. В разі необхідності ми можемо легко збільшувати та зменшувати обчислювальну потужність за декілька хвилин.

### 3.2 Використання даних

Для супутникового моніторингу посівів нам в першу чергу нам потрібні часові серії супутникових знімків, та дані навчальних, валідаційних та тестових наборів. Так як основною проблемою моніторингу є захмарені знімки, тому розробивши власні методи, які є наближені до світових, ми створюємо безмарні композити. Для покращення результату безхмарних комополітів, використовується попередня обробка геоданих, а саме вирізання хмар та заміна значень за найближчі дні.

Зазвичай супутникові знімки зберігаються на платформі коперникуса[26]. Цей сервіс пропонує нам безкоштовно супутникові дані будь якої території за будь який часовий період. Але щоб працювати з цими даними їх потрібно завантажувати і це значно сповільнює роботу, адже супутникові геодані займають чимало пам'яті. Тому для прискорення цього процесу використовується хмарне середовище, Amazon, де за допомогою цих сервісів ми можемо геодані Sentinel-2 легко завантажити в хмару.

#### 3.2.1 Векторні дані

Векторні дані представляють собою таку категорію даних, яка з'являється у знімках, де кожен елемент представлений своєрідною геометрією, або вектором. Така інформація є дуже корисною коли мета задачі полягає в обробці чи дослідження геометричних просторів.

Для того щоб створити класифікацію на потрібні подати на вхід дані. Так як в даній системі використовується навчання з вчителем, на потрібно підготувати векторні дані, за допомогою яких в подальшому буде проводитися класифікація. Для отримання більш чітких результатів, дуже важливий

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		36

правильний підбір даних. Тому дані які використовуються в даній системі були підготовленні Інститутом космічних досліджень. У векторні дані які використовуються для класифікації входять наземні вимірювання, полігони з кадастру та інші векторні дані.

Для подальшої взаємодії з даними нам потрібно привести їх до спільного масштабу, адже звичайного завантаження полігонів недостатньо. Кожен полігон маж взаємодіяти між растровими положеннями піксельними та лінійними координатами на карті.

### 3.3 Розробка функціоналу системи

Для того, щоб розробити систему попередньої обробки даних, було вибрано мову програмування високого рівня Python. Високорівнева мова програмування Python була створена нідерландським програмістом Гвідо Ван Россумом. На даний час все більше розробників переходять на цей безкоштовний продукт по створенню якісних та ємних програм, щоб спростити процес програмування. Дружелюбний інтерфейс мови та універсальність застосування роблять Python однією із потужних технологій розробки програмного забезпечення та відкриває новий спектри можливостей в плані підходу та швидкості задач. Зараз, для створення сучасних програм використовують не тільки вбудований функціонал мови, а й активно користуються додатковими бібліотеками, без яких зараз вже і складно уявити процес розробки використовуючи даний інструмент. Дана мова вже зараз дає змогу оброблювати великі об'єми інформації, створювати складні та сучасні сайти, використовуватись у науковій та аналізу даних.

У роботі були використані такі прикладні бібліотеки Python як: Numpy, scikit-learn, gdal, rasterio, scipy, xarray та pandas.

Python є дуже потужною мовою, щоб задовільнити потреби в обробці геоданих. Цю мову можна використовувати не тільки для обробки та візуалізації даних, але й для адміністрування системи, веб-розробки, програмування баз даних. Зараз для python розробляється дуже багато різних

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		37

бібліотек для вирішення різного типу задач, але навіть якщо для вирішення тієї чи іншої проблеми не достатньо python, то можна вирішити ці проблеми за допомогою інтеграції з іншими мовами.

За допомогою цієї мови в поєднанні з open data cube в даній архітектурі здійснюється попередня обробка даних, а саме композитування та вирізання хмар.

Також за допомогою цієї мови виконується машинне навчання, адже для цієї мови розроблена бібліотека яка називається scikit-learn. Ця бібліотека пропонує нам вбудовані методи машинного навчання з глибокими налаштуваннями. Для отримання результатів нам потрібно подати на вхід наші векторні дані, які за допомогою методів машинного навчання будуть використанні для тренування, подальшої класифікації та валідації.

Дана додаткова прикладна бібліотека була використана як інструмент тестування методів машинного навчання, оскільки вона володіє дуже великим набором необхідних застосунків у сфері аналізу даних та математичних методів штучного інтелекту. Також за допомогою неї проводилась робота по оцінці точності алгоритмів.

Для обробки даних була обрана програма jupyter notebook. Це дуже зручна програма та надзвичайно потужний інструмент для розробки різних проектів в області науки даних. Інтерфейс програми є зручним, для того щоб почати працювати нам достатньо запустити цю програму, та почати писати код. На відміну від інших середовищ розробки програмного забезпечення в jupyter notebook можна запускати код шматками. Тобто ми окремо запускати кожну функцію, що значно полегшує та пришвидшує обробку даних.

### 3.3.1 Розробка uml діаграми

Використовуючи хмарну систему open data cube користувач може задати параметри, або змінити їх, створювати хмарні та безхмарні композити та будувати карту. Функціонал попередньої обробки даних представлений на діаграмі прецедентів на рисунку 3.2.

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		38

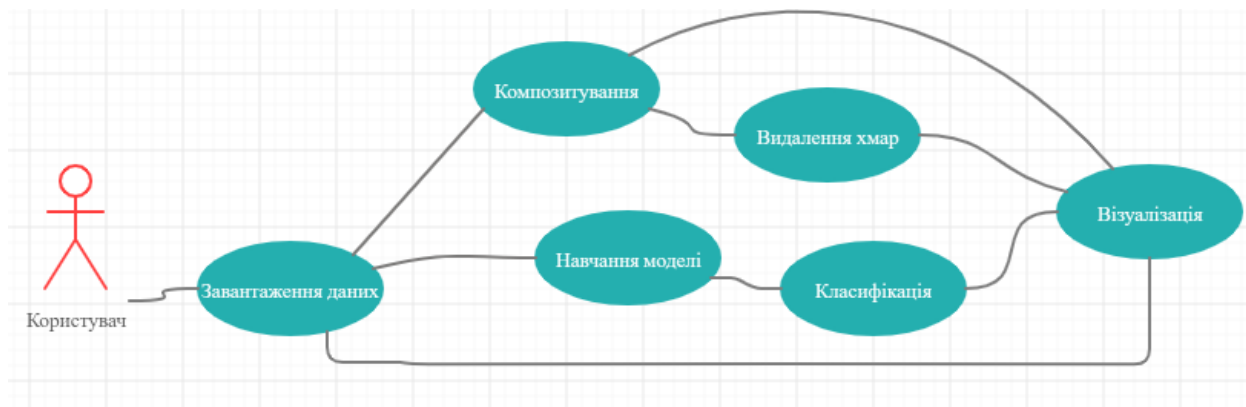


Рис. 3.2 Діаграма прецедентів

Користувач у даній системі буде мати можливість виконувати наступні дії:

- Завантаження потрібних даних використовуючи програмне забезпечення open data cube.
- Створювати композити
- Використовуючи створені композити видаляти хмари
- Навчати модель за власною вибіркою
- Створювати карту класифікації
- Візуалізовувати всі дані

### 3.4 Функціонал попередньої обробки даних

Для полегшення роботи з open data cube було розроблені власні класи методи яких допомагають полегшити роботу з обробкою даних, також розроблені методи полегшують взаємодію між існуючими бібліотеками. Діаграма класів наведена в додатку 3 та принципова схема роботи наведена в додатку 4.

Для розробки функціоналу попередньої обробки була використана бібліотека GDAL. Даний програмний застосунок представляє собою систему вкрай важливих програм, за допомогою яких відбувається дослідження та аналіз геопросторових даних. Дана потужна бібліотека дає можливість використовувати її на різних мовах програмування, що робить її більш

універсальним інструментом для розробки програм. В роботі використовувались функції імпортування файлів супутникових знімків, обробки та правильне зберігання у файл Geotiff.

Також хорошим доповненням цієї бібліотеки є rasterio. Rasterio є осучасненою та більш модифікованою версією gdal, хоча саме вона його використовує як базовий рушій. Дана бібліотека також орієнтована на роботу, дослідження та обробку супутникових даних, що дозволяє швидко та легко керувати знімками. Дана технологія була використана в роботі у вигляді компоненти зберігання композитів із правильною прив'язкою координат та інших мета дани.

За основу розробки були взяті дані sentinel-2 які зображені на рисунку 3.3 ці дані знаходяться на open data cube. На основі цих даних буде виконуватись попередня обробка супутникових знімків, а саме створення хмарних та безхмарних композитів.

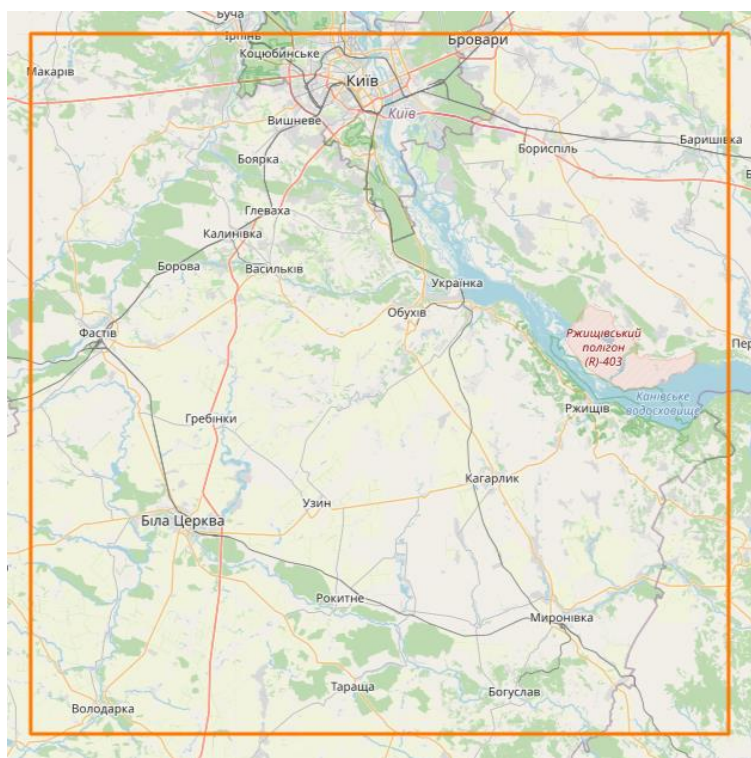


Рис. 3.3 карта яка використовуються для попередньої обробки

Даний датасет містить дані за 2019 рік, а саме з 3 березня, по 28 вересня 2019 року. Для того щоб зручно користуватися даними був створений клас data. За допомогою цього класу нам значно легше користуватися даними.

Всі дані зберігаються в масиві xarray. За допомогою даної бібліотеки можна легко керувати даними з часовим простором чи іншою додатковою віссю. Дана програма активно використовується для дослідження великих даних, таких як масивів замірів температури, днями та координатами. У роботі використовувалась для завантаження датасетів та отримання з них необхідної інформації. Xarray представляє мітки у вигляді розмірів, координат та атрибутів поверх необроблених масивів, схожих на NumPy, що дозволяє отримати більш простіший та більш стислий та менш схильний до помилок код. За допомогою xarray даними легко керувати та візуалізувати їх. Нижче ми можемо бачити як виглядає весь датасет в xarray масиві за весь час.

```
<xarray.Dataset>
Dimensions:  (time: 128, x: 11840, y: 11825)
Coordinates:
  * time (time) datetime64[ns] 2019-03-02T09:15:55.500000 ... 20
19-09-28T09:12:48
  * y (y) float64 5.6e+06 5.6e+06 5.6e+06 ... 5.482e+06 5.482e+0
6
  * x (x) float64 2.653e+05 2.653e+05 2.653e+05 ... 3.837e+05 3.
837e+05
Data variables:
  red(time, y, x) int16 dask.array<shape=(128, 11825, 11840),
chunksize=(1, 200, 2000)>
  green(time, y, x) int16 dask.array<shape=(128, 11825, 11840)
, chunksize=(1, 200, 2000)>
  blue(time, y, x) int16 dask.array<shape=(128, 11825, 11840),
chunksize=(1, 200, 2000)>
  nir(time, y, x) int16 dask.array<shape=(128, 11825, 11840),
chunksize=(1, 200, 2000)>
  slc(time, y, x) uint8 dask.array<shape=(128, 11825, 11840),
chunksize=(1, 200, 2000)>
Attributes:
  crs:  EPSG:32636
```

Цей датасет включає в себе розмірність x,y кількість днів які містяться у всьому датасеті, та попередньо вибрані канали такі як red, green, blue, nir, slc для роботи з цим датасетом, та містить атрибути загального стандарту EPSG:32636

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		41

### 3.4.1 Обробка всього датасету

Для того щоб за один раз обробити весь датасет потрібно використовувати багато обчислювальних ресурсів. Тому для оптимізації було створено клас `DataLoader` за допомогою якого ви розбиваємо весь датасет на задані нам шматки. Це значно пришвидшує вичитку даних окремого шматка. Для прикладу розбивши всю територію на шматки 1000/1000 ми отримуємо дану матрицю:

```
from scripts.data import DataLoader
dc = datacube.Datacube(app = 'my_app', config = '/home/localuser/.datacube.conf')
data = DataLoader(dc = dc,
                  latitude = latitude,
                  longitude = longitude,
                  time = time,
                  chunk_size = (1000,1000),
                  view_with_nodata = True)
data.load_data(True)
```

Matrix of chunks:

```
[[ 0  1  2  3  4  5  6  7  8  9 10 11]
 [12 13 14 15 16 17 18 19 20 21 22 23]
 [24 25 26 27 28 29 30 31 32 33 34 35]
 [36 37 38 39 40 41 42 43 44 45 46 47]
 [48 49 50 51 52 53 54 55 56 57 58 59]
 [60 61 62 63 64 65 66 67 68 69 70 71]
 [72 73 74 75 76 77 78 79 80 81 82 83]
 [84 85 86 87 88 89 90 91 92 93 94 95]
 [96 97 98 99 100 101 102 103 104 105 106 107]
 [108 109 110 111 112 113 114 115 116 117 118 119]
 [120 121 122 123 124 125 126 127 128 129 130 131]
 [132 133 134 135 136 137 138 139 140 141 142 143]]
```

Рис. 3.4 Застосування методу класу `DataLoader` який розбивє датасет на шматки

Створена матриця із шматків 1000/1000 пікселів дозволяє нам легко ітеруватися по шматкам, це значно прискорює роботу. Раніше, щоб досягнути до певного регіону, раніше нам потрібно було задавати довготу та широту.

Також розроблений метод в цьому класі дозволяє не завантажувати шматок в пам'ять, а просто індексуватися по ним.



### 3.5 Візуалізація геоданих

Щоб не завантажувати кожного разу всі дані, адже кожен композит займає близько 500мб, використовується сервіс для зберігання інформації AWS S3. За допомогою цього сервісу наші дані є захищеними та до них легко доступитися в будь який момент часу, також сервіс має зручний інтерфейс, та вбудовані програми з глибокими налаштуваннями для користувачів.

Так як в opendatacube немає інструментів за допомогою яких можна було б візуалізувати дані, тому було розроблено клас view який базується на бібліотеці matplotlib. За допомогою методів цього класу ми можемо візуалізувати будь який знімок в форматі numpy array та хarray, та маску хмарності. Також для більш детальної візуалізації супутникових знімків було використано безкоштовне програмне забезпечення Quantum GIS (QGIS).

Також розроблені власні методи дозволяють нам отримати всю інформацію про даний датасет, адже сама система sentinel-2 нам не надає дану інформацію. Метод get\_info класу data дозволяє нам отримати інформацію яка полегшить нам подальшу обробку. За допомогою цього методу ми можемо бачити на скільки кожен знімок є захмареним, або чи містить знімок дані.

На рис 3.4 ми можемо бачити дані одного шматка 500/500 пікселів за серпень 2019 року. Дані містять відсоток хмарності, відсоток димки, та nodata.

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		43

	time	cloudiness_percentage	cirrus_percentage	nan_percentage of whole comp
0	2019-08-02T09:27:22.000000000	0.000000	0.000000	1.0
1	2019-08-04T09:15:46.000000000	0.516938	0.295000	0.0
2	2019-08-06T09:05:03.000000000	0.017785	0.004081	0.0
3	2019-08-07T09:25:35.000000000	0.000000	0.000000	1.0
4	2019-08-09T09:11:58.000000000	0.047918	0.045590	0.0
5	2019-08-12T09:27:35.000000000	0.000000	0.000000	1.0
6	2019-08-13T08:54:30.000000000	0.000000	0.000000	1.0
7	2019-08-14T09:18:27.500000000	0.332068	0.000000	0.0
8	2019-08-18T08:56:14.000000000	0.000000	0.000000	1.0
9	2019-08-19T09:18:53.000000000	0.000000	0.000000	1.0
10	2019-08-21T09:07:34.000000000	0.001953	0.000000	0.0
11	2019-08-24T09:12:09.500000000	0.000000	0.000000	1.0
12	2019-08-26T09:01:55.500000000	0.304218	0.000000	0.0
13	2019-08-27T09:24:28.000000000	0.000000	0.000000	1.0
14	2019-08-28T08:54:45.000000000	0.000000	0.000000	1.0
15	2019-08-29T09:18:52.000000000	0.002272	0.000000	0.0
16	2019-08-31T09:02:47.500000000	0.038970	0.000000	0.0

Рис. 3.4 Приклад роботи функції get\_info

Дана таблиця побудована за допомогою бібліотеки pandas, що дозволяє нам з легкістю візуалізувати дані. Pandas є дуже відомою бібліотекою, яка являє собою середовище для візуалізації та первинною обробкою даних. Дана технологія використовувалась для відображення чисельної інформації і характеристик про вхідний датасет.

Використовуючи ці дані ми відсортовуємо знімки які містять більше 70% хмарності. Знімки які мають невелику кількість даних не відкидаються, адже ці дані є потрібними, наприклад для заміни значень за найближчі дні. Зазвичай в датасеті інколи трапляються знімки які не є потрібними, для аналізу, тобто знімок немає даних, або 70-80% знімку є захмареним. Тому щоб пришвидшити обробку геоданих, було створено функції за допомогою яких можна прибирати непотрібні знімки, та відкидати знімки в яких хмарність є більшою за задану.

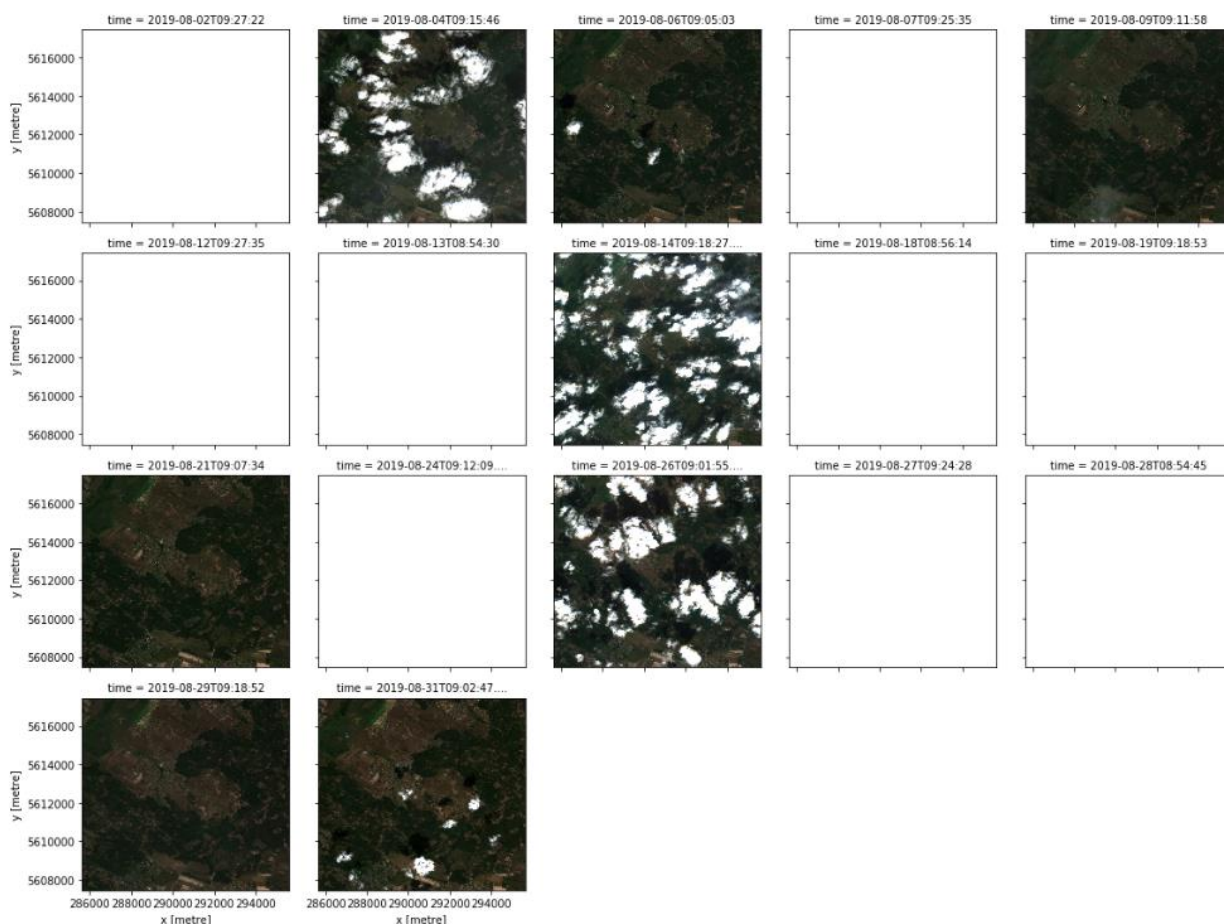


Рис. 3.5 Візуалізація знімків без даних та захмарених

Після застосування власного методу, який нам дозволяє відкидати непотрібні дані, ми можемо бачити як змінилася таблиця з даними.

	time	cloudiness_percentage	cirrus_percetage	nan_percentage of whole comp
0	2019-08-06T09:05:03.000000000	0.017785	0.004081	0.0
1	2019-08-09T09:11:58.000000000	0.047918	0.045590	0.0
2	2019-08-14T09:18:27.500000000	0.332068	0.000000	0.0
3	2019-08-21T09:07:34.000000000	0.001953	0.000000	0.0
4	2019-08-26T09:01:55.500000000	0.304218	0.000000	0.0
5	2019-08-29T09:18:52.000000000	0.002272	0.000000	0.0
6	2019-08-31T09:02:47.500000000	0.038970	0.000000	0.0

Рис. 3.6 Результат відсортування

Та також можна побачити як це змінилося візуально на рисунку 3.7.

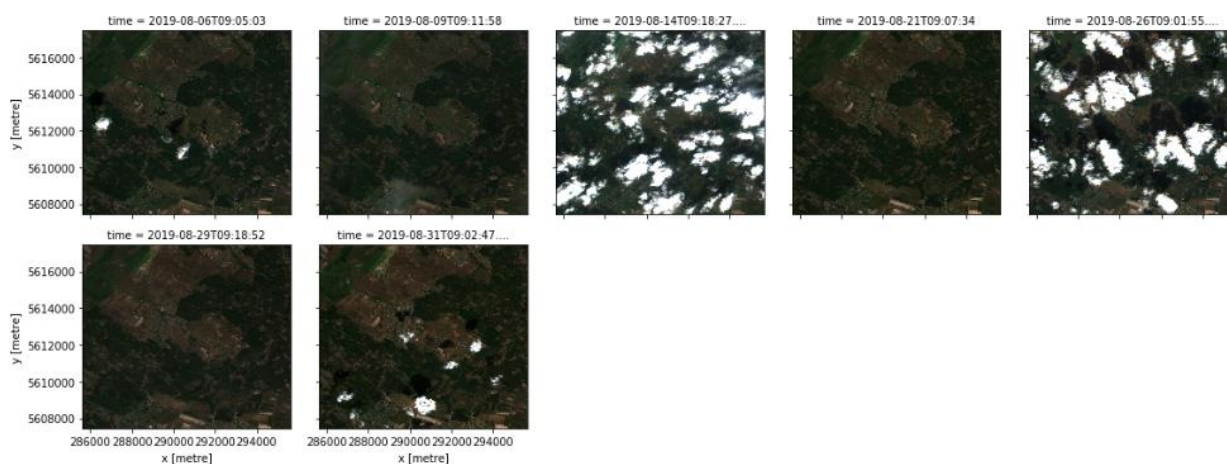


Рис. 3.7 Візуалізація знімків після відсортування

На рисунку 3.8 створено композит в RGB каналах, тобто кольоровий композит, який є результатом склейки трьох мультиспектральних каналів, але як ми можемо бачити цей знімок є захмареним.

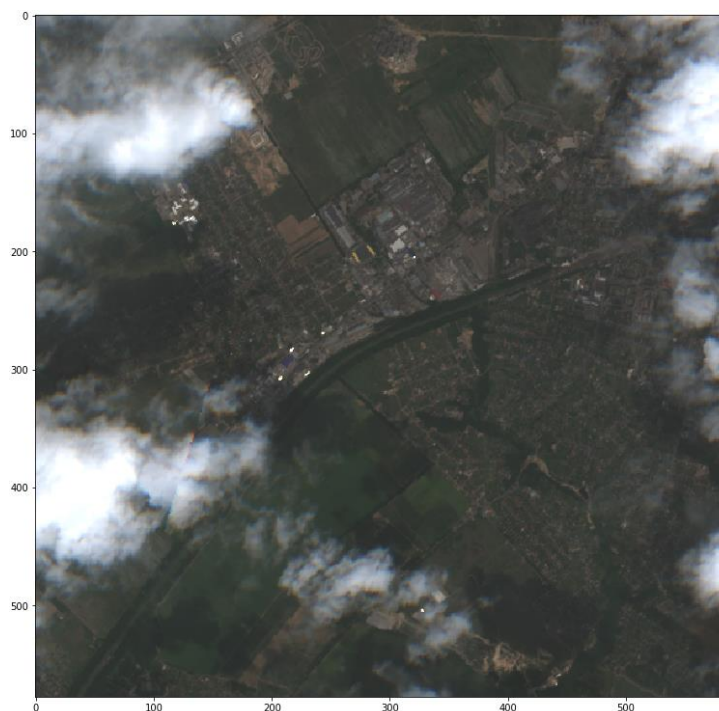


Рис. 3.8 Хмарний композит

Використовуючи маску класифікації, яку пропонує нам sentinel-2, бібліотеку scipy та алгоритм Dilation, який значно покращує результати, способом розширення маски, ми можемо бачити результат бехмарного композиту.

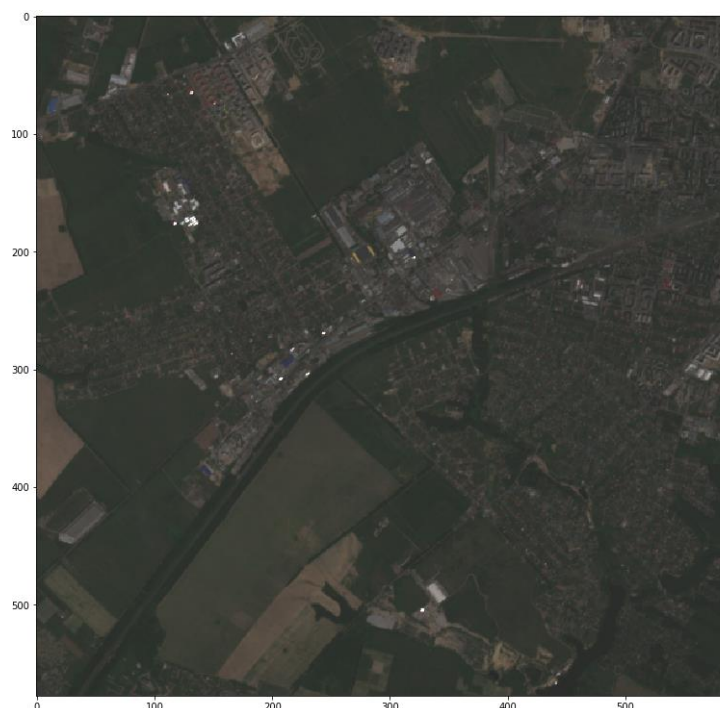


Рис.3.9 Безхмарний композит

Scipy надає можливість використовувати математичний і статистичний апарат у мові програмування Python. Майже всі статистики та математичні оператори чи алгоритми для обробки даних чи перевірки гіпотез. Даний інструмент використовувався у відомі сфері математиці як морфологія. І саме завдяки ньому обраховувались оцінки схожості чи точності алгоритмів.

### 3.6 Завантаження даних

Інколи за допомогою класу який був розроблений для візуалізації складно побачити недоліки які є на знімку. Для кращої візуалізації даних використовувалась програма (Quantum GIS). Але для того щоб подивитися знімки в цій програмі нам потрібно було їх завантажити. Створений клас driver зі своїми методами дозволяє нам завантажувати зображення. За допомогою бібліотек rasterio та gdal було розроблено методи класу які дозволяють нам зберігати зображення як з хатгау так і з numpy array, зберігаючи метадані, адже зберігання метаданих є дуже важливою частиною. В метадах зберігаються всі дані щодо знімку, а саме проекції, кількість каналів, стандартизація карти.

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		47

Також методи цього класу дозволяють завантажувати зображення з хмарними композитами так і безхмарними.

					ІАЛЦ.467200.002 ПЗ	Арк.
						48
Зм.	Арк.	№ докум.	Підпис	Дата		

### ВИСНОВКИ ДО РОЗДІЛУ 3

Супутникові дані потребують великої обчислювальної потужності, адже завантаження знімків та попередня обробка даних, є важливим моментом для подальшого аналізу. Тому для вирішення цієї проблеми було розроблено функціональну, структурну та класову діаграму запропонованої інформаційної системи, на основі компонентів хмарної платформи Amazon Web Services. Адже основними перевагами хмарного сервісу є швидкість завантаження даних, потужні обчислювальні ресурси, та велика ємність зберігання даних.

Отже, використовуючи хмарні сервіси ми можемо значно скоротити швидкість обчислення даних, що значно прискорить отримання результатів.

					ІАЛЦ.467200.002 ПЗ	Арк.
						49
Зм.	Арк.	№ докум.	Підпис	Дата		

## РОЗДІЛ 4

### ЕКСПЕРЕМЕНТИ ТА ТЕСТУВАННЯ

#### 4.1 Швидкість роботи

В даному розділі було протестована попередня обробка даних на локальному комп'ютері та в хмарі.

Для перевірки швидкості на локальному комп'ютері було використано:

- Комп'ютер на базі процесора Intel i-7 4770MQ 2,7 ГГц
- Оперативна пам'ять 6 Гб.
- Вільний простір жорсткого диску 1.5ГбГГц

Для перевірки швидкості на сервері було використано:

- Процесор Intel Xeon Platinum 8175M 2.5
- Оперативна пам'ять 32Гб
- SSD

Для того, щоб виконати попередню обробку даних нам потрібно було викачати знімки за цілий місяць з сайту коперникуса[13],адже щоб отримати безмарні композити нам потрібно замінити вирізані пікселі за найближчі дати. Виконавши завантаження було створено композити, та застосовано алгоритм dilation.

- Час обробки знімку на локальному комп'ютері становить 14 хвилин та 43 секунди.
- Виконавши цю попередню обробку даних хмарі, час обробки становить 34секунди.

Це означає, що в хмарі набагато краще проводити дослідження, та головною перевагою є те що дані не потрібно завантажувати, адже завдяки сервісам AMAZON вони знаходяться в хмарі, що значно прискорює роботу.



## 4.2 Уточнення маски хмари

Для покращення результатів, було використано алгоритм dilation для рисунку 4.1.

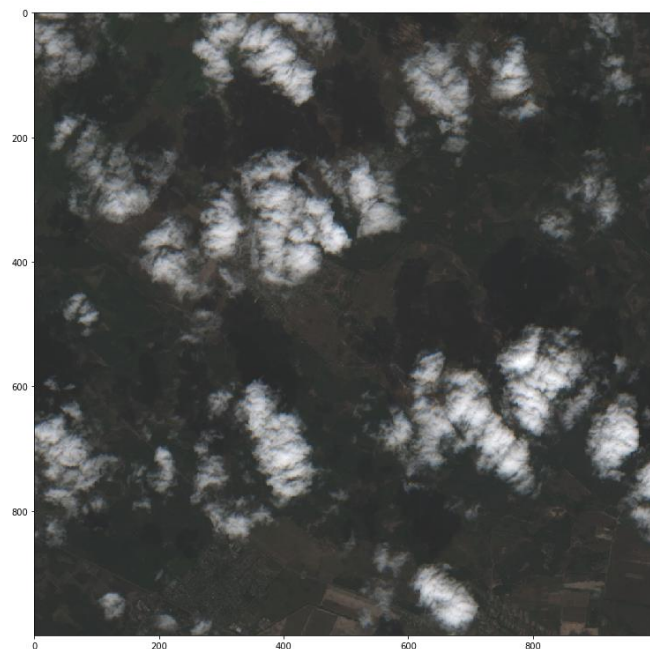


Рис. 4.1 Захмарений знімок з розмірністю (1000/1000 пікселів)

Створивши відповідні функції за допомогою яких ми можемо бачити маску хмар яка покриває весь знімок на рисунку 4.2.

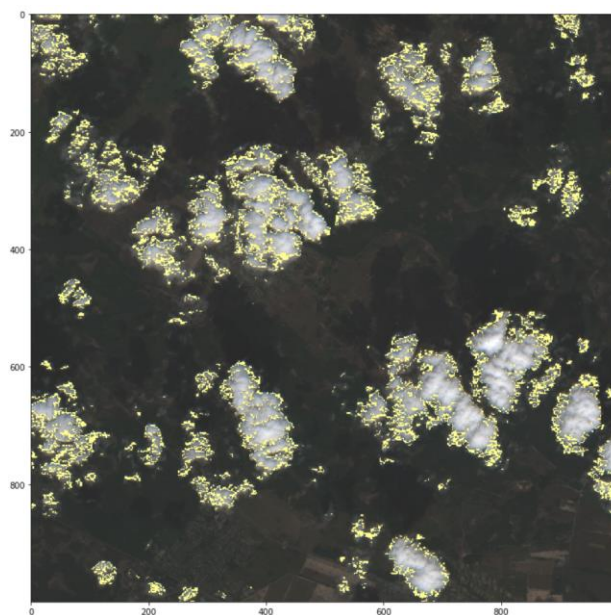


Рис. 4.2. Маска хмар знімку наведеного вище

Створивши відповідну функцію за допомогою якої ми можемо порахувати відсоток хмар на знімку, ми можемо бачити на скільки збільшилась маска при розширенні. Оригінальна маска хмар на композиті становить 24%. Використавши алгоритм dilation ми збільшили маску на 6% відсотків і на даний момент вона становить 30%, тому це означає що маска тепер охоплює значно більше хмар, а саме краї.

На рисунку 4.3. ми можемо бачити готовий безхмарний композит з застосуванням алгоритму dilation.

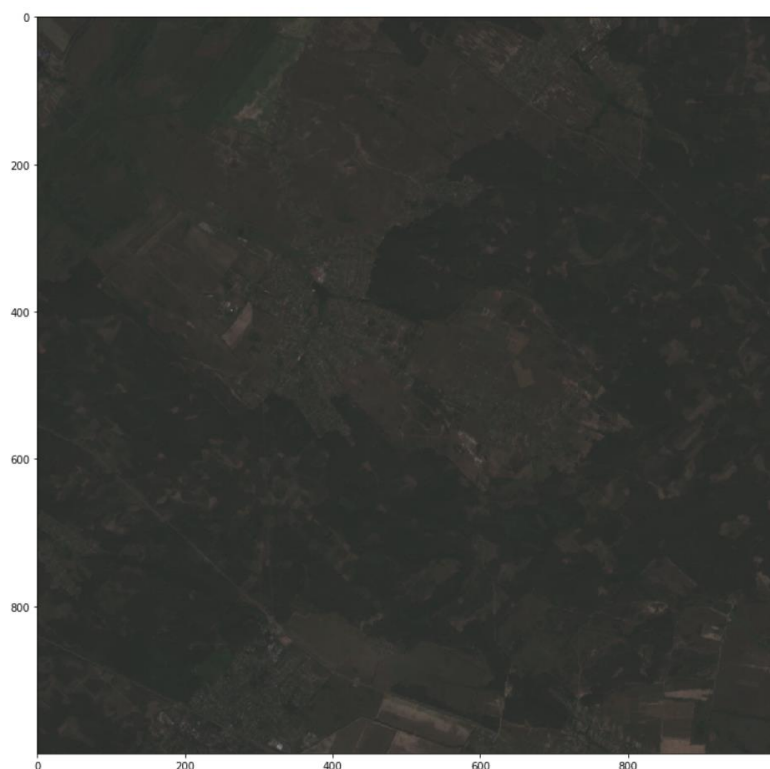


Рис 4.3. Безхмарний композит із застосуванням алгоритму dilaton

Даний композит є результатом попередньої обробки, та може використовуватися для подальшого аналізу.

#### 4.3NDVI

Для того щоб перевірити індекс вегетації за допомогою створеної функції за формулою, було візуалізовано NDVI для хмарного та безхмарного композиту.

					ІАЛЦ.467200.002 ПЗ	Арк.
						52
Зм.	Арк.	№ докум.	Підпис	Дата		

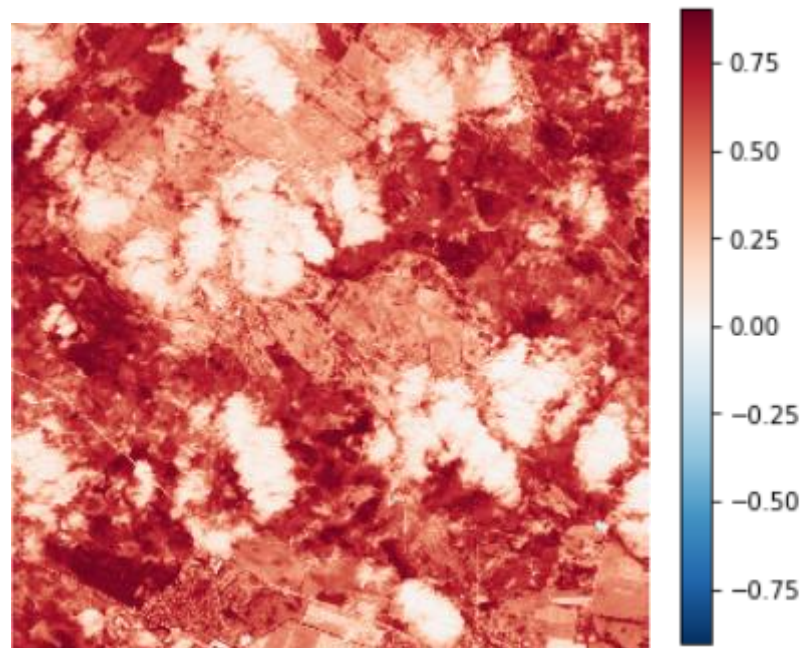


Рис.4.4 NDVI для хмарного композиту

На рисунку видно, що більшість хмар закривають більшість знімку, тому для вирішення цієї проблеми було створений безхмарний композит за власним алгоритмом.

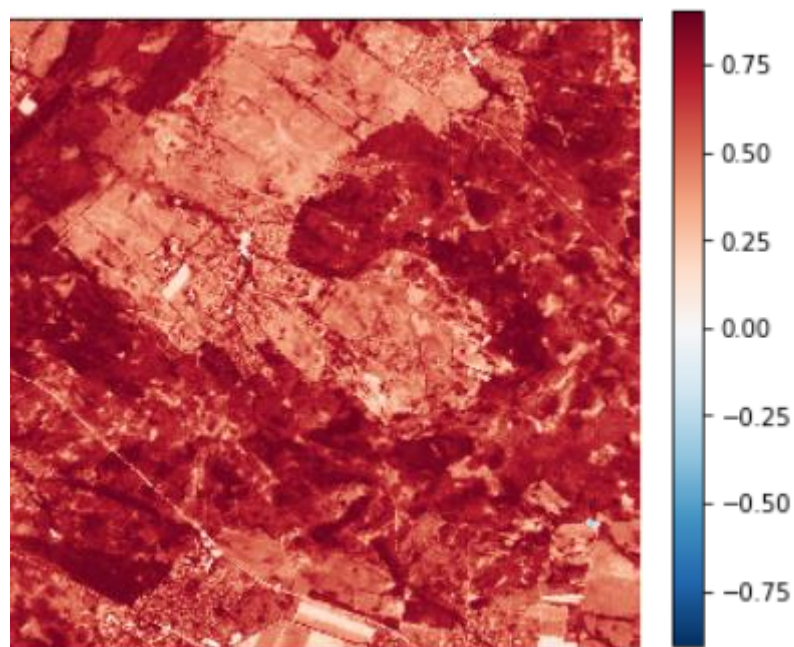


Рис.4.5 NDVI для безхмарного композиту.

Завдяки мультиспектральним каналам ми можемо бачити як дана шкала нам показує вегетаційний індекс, для більш здоровіших насаджень колір є

більш інтенсивним, тому густина насаджень є більшою, а там де низька фотосинтетична активність колір є більш світлішим, це демонструє нам, що дана область є проблемною. Для порівняння можна побачити цей знімок для якого застосовувався NDVI в форматі RGB.

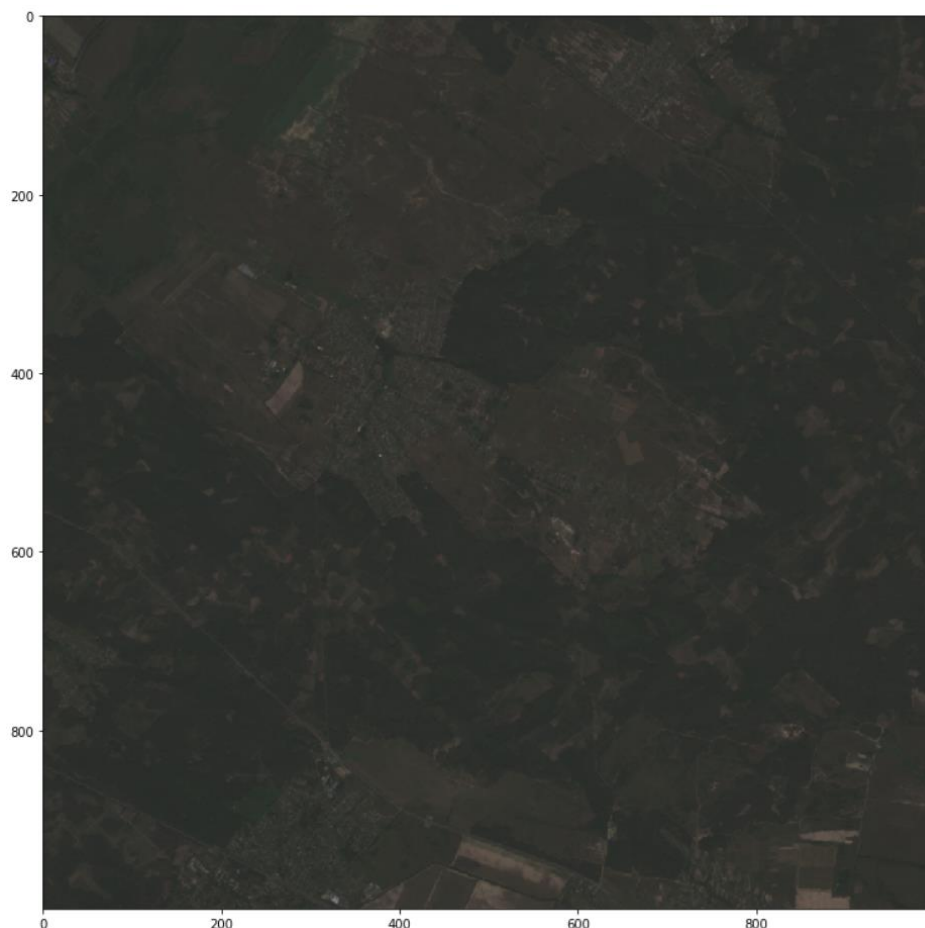


Рис. 4.6 Безмарний композит без застосування NDVI

Ми можемо порівняти цей знімок з прикладом використання NDVI. Де область є більш зеленішою, на рисунку NDVI область є більш інтенсивнішою, та навпаки.

#### 4.4 Класифікація

На основі даних які підлягали попередній обробці була виконана класифікація за допомогою алгоритму випадкового лісу, та багатошарового перцептрону. За допомогою цих методів машинного навчання була

					ІАЛЦ.467200.002 ПЗ	Арк.
						54
Зм.	Арк.	№ докум.	Підпис	Дата		



побудована карта, та порахована точність класифікації для кожного методу. Для карти класифікації використовувались дані за липень 2019 року.

Карта побудована багат шаровим перцептроном має точність 0.74. На рисунку 4.7 ми можемо бачити карту класифікації



Рис. 4.7 Карта класифікації MLP

Також було побудовано карту класифікації за допомогою random forest. Точність цієї карти складає 0.59. Карта класифікації зображена на рисунку 4.8.



Рис. 4.8 Карта класифікації за допомогою класифікатора random forest

Проаналізувавши дані результати ми можемо бачити що точність багат шарового перцептрону є набагато більшою. Для тестування також було заміряно час. Замірювання часу використовувалося для системи яка знаходиться в хмарі, на локальному комп'ютері заміряти час нажаль не вдалося, адже система потребує багато ресурсів.

Час класифікації для MLP:

- 37 хвилин.

Час класифікації для RF:

- 4 хвилини.

Random Forest працює значно швидше, але демонструє нижчий рівень точності ніж MLP.

					ІАЛЦ.467200.002 ПЗ	Арк.
						56
Зм.	Арк.	№ докум.	Підпис	Дата		

## ВИСНОВКИ ДО РОЗДІЛУ 4

В даному розділі було апробовано працездатність модулів на прикладі набору геопросторових даних для території Київської області. Так як система сільськогосподарського моніторингу потребує великих об'ємів ресурсів, тому протестувавши та проаналізувавши алгоритми які використовуються для попередньої обробки даних, було помічено, що інколи обробку даних неможливо застосувати на локальному комп'ютері, тому можна сказати, що краще виконувати всі обчислення в хмарі, адже це є більш ефективним методом та значно пришвидшує роботу.

					ІАЛЦ.467200.002 ПЗ	Арк.
						57
Зм.	Арк.	№ докум.	Підпис	Дата		

## ВИСНОВКИ

В світі є безліч проектів та систем які займаються супутниковим моніторингом, в тому числі й для України. Так як для України ще не реалізовано системи моніторингу сільськогосподарських посівів для високого просторового розрізнення, тому було розроблено власні методи які є схожими до світових на основі компонентів хмарної платформи Amazon Web Services та безкоштовного програмного забезпечення DataCube.

В даній роботі було розроблено алгоритми, які дозволяють розширити маску хмар, створювати хмарні та безхмарні композити, які є основою модуля попередньої обробки.

Модуль попередньої обробки є частиною всієї хмарної системи моніторингу сільськогосподарських посівів, яка складається з самого модуля попередньої обробки, модуля підготовки даних та машинного навчання. Це дає можливість розв'язувати задачі для хмарної системи супутникового моніторингу посівів використовуючи часові ряди.

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		58



## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Remote sensing-based global crop monitoring: Experiences with China's CropWatch system [Стаття] / B.Wu, J. Meng, Q. Li, N. Yan // International Journal of Digital Earth. – 2014. – Режим доступу до ресурсу: [https://www.researchgate.net/publication/261016698\\_Remote\\_sensing-based\\_global\\_crop\\_monitoring\\_Experiences\\_with\\_China's\\_CropWatch\\_system](https://www.researchgate.net/publication/261016698_Remote_sensing-based_global_crop_monitoring_Experiences_with_China's_CropWatch_system).
2. CropWatch [Електронний ресурс] – Режим доступу до ресурсу: <http://www.cropwatch.com.cn/htm/en/methodology.shtml>.
3. CropMonitor [Електронний ресурс] // NASA – Режим доступу до ресурсу: <https://nasaharvest.org/project/geoglam-crop-monitor-amis>
4. The GEO Global Agricultural Monitoring (GEOGLAM) Initiative [Електронний ресурс] // NASA – Режим доступу до ресурсу: <https://nasaharvest.org/index.php/project/geo-global-agricultural-monitoring-geoglam-initiative>.
5. MARS [Електронний ресурс] // EU SCIENCE HUB – Режим доступу до ресурсу: <https://ec.europa.eu/jrc/en/mars>.
6. MARS Agricultural monitoring [Електронний ресурс] // EU SCIENCE HUB – Режим доступу до ресурсу: <https://ec.europa.eu/jrc/en/research-topic/agricultural-monitoring>
7. Open Data Cube [Електронний ресурс] – Режим доступу до ресурсу: <https://www.opendatacube.org/>
8. Документація Open Data Cube [Електронний ресурс] – Режим доступу до ресурсу: <https://readthedocs.org/projects/opendatacube/downloads/pdf/latest/>.
9. What is the Open Data Cube? Cube [Електронний ресурс] – Режим доступу до ресурсу: <https://medium.com/opendatacube/what-is-open-data-cube-805af60820d7>
10. Overview of the Open Data Cube Initiative [Стаття] / [Brian Killough] //

Режим доступу до ресурсу:

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		59

[https://www.researchgate.net/publication/328995502\\_Overview\\_of\\_the\\_Open\\_Data\\_Cube\\_Initiative](https://www.researchgate.net/publication/328995502_Overview_of_the_Open_Data_Cube_Initiative)

- 11.ESA [Електронний ресурс] Режим доступу до ресурсу:  
<http://www.esa.int/>
- 12.Sentinel-2 [Електронний ресурс] // ESA – Режим доступу до ресурсу:  
<https://sentinel.esa.int/web/sentinel/missions/sentinel-2>.
13. Sentinel-2 [Електронний ресурс] // ESA – Режим доступу до ресурсу:  
<https://sentinel.esa.int/web/sentinel/missions/sentinel-2/overview>
- 14.Methods for Creating Cloud Free Satellite Imagery: Sentinel 2 [Електронний ресурс] // ESA – Режим доступу до ресурсу:  
<https://www.gislounge.com/methods-creating-cloud-free-satellite-imagery-sentinel-2/>
- 15.Scene Classification [Електронний ресурс] // ESA – Режим доступу до ресурсу:  
<https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm>
- 16.Spectral bands for the SENTINEL-2 sensors (S2A & S2B) [Електронний ресурс] // ESA – Режим доступу до ресурсу:  
<https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument>.
- 17.Facts and figures [Електронний ресурс] // ESA – Режим доступу до ресурсу:  
<https://directory.eoportal.org/web/eoportal/satellite-missions/c-missions/copernicus-sentinel-2>
- 18.Random forest [Електронний ресурс] – Режим доступу до ресурсу:  
[https://uk.wikipedia.org/wiki/Random\\_forest](https://uk.wikipedia.org/wiki/Random_forest).
- 19.Многослойный перцептрон Румельхарта [Електронний ресурс] – Режим доступу до ресурсу: [https://en.wikipedia.org/wiki/Multilayer\\_perceptron](https://en.wikipedia.org/wiki/Multilayer_perceptron)
- 20.Введение в нейронные сети [Статья] – Режим доступу до ресурсу:  
<http://blog.skahin.ru/2018/06/blog-post.html>

					ІАЛЦ.467200.002 ПЗ	Арк.
Зм.	Арк.	№ докум.	Підпис	Дата		60

21. Image Composites [Електронний ресурс] – Режим доступу до ресурсу:  
[http://gsp.humboldt.edu/OLM/Courses/GSP\\_216\\_Online/lesson3-1/composites.html](http://gsp.humboldt.edu/OLM/Courses/GSP_216_Online/lesson3-1/composites.html)
22. Classification Mask Generation [Електронний ресурс] // ESA – Режим доступу до ресурсу: <https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm>.
23. M. Goyal// Morphological Image Processing//International Journal of Computer Science and Technology // IJCST Vol. 2, Issue 4, Oct.—2011 P. 161-165.
24. CLOUD APPROACH TO AUTOMATED CROP CLASSIFICATION USING SENTINEL-1 IMAGERY [Електронний ресурс] / [A. Shelestov, M. Lavreniuk, N. Kussul та ін.] // EOSDA, Kyiv, Ukraine
25. Yailymov, Bohdan; Kussul, Nataliia; Shumilo, Leonid; Lavreniuk, Mykola; Shelestov, Andrii; Yailymova, Hanna; Bilokonska, Yuliia; Kalinchuk, Anastasia; Parkhomchuk, Olexandr; Dikhtyar, Ivan, "SDG Indicators 15.1.1 And 15.3.1 Assessment Using Open Data Cube Technology", In EO for Agriculture under Pressure 2020. Submitted
26. Copernicus Open Access Hub [Електронний ресурс] // Copernicus – Режим доступу до ресурсу: <https://scihub.copernicus.eu/dhus/#/home>.

# ДОДАТОК 1

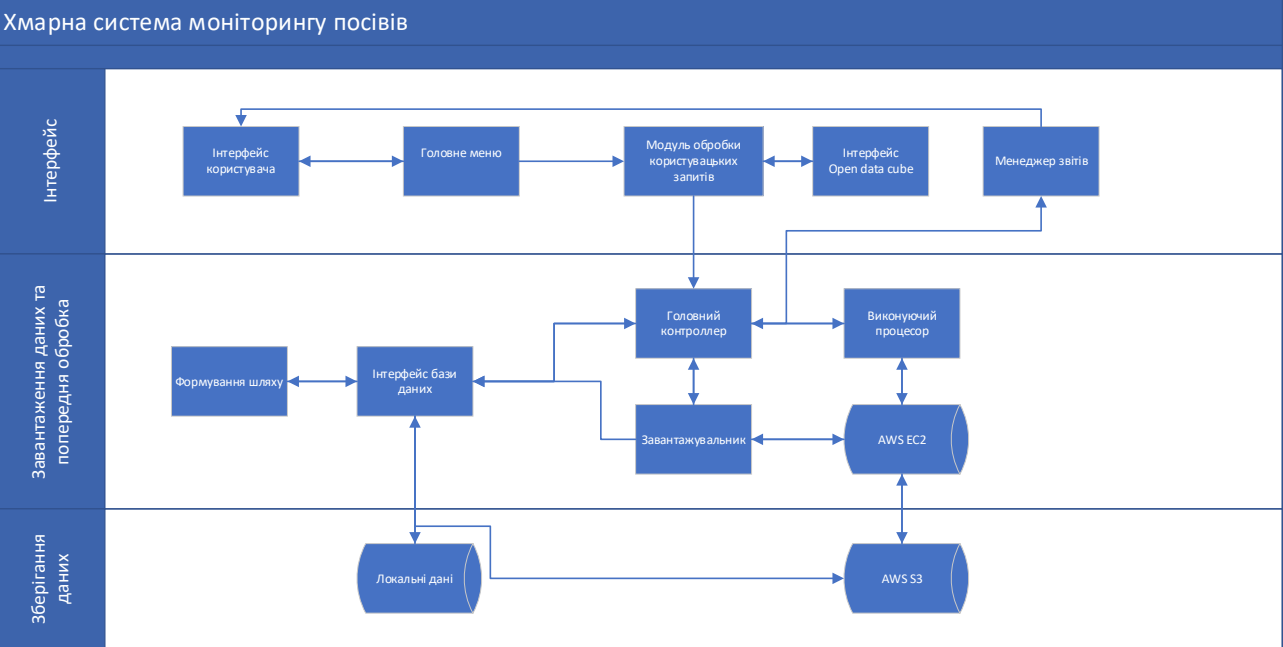
Хмарна система супутникового моніторингу посівів

ФУНКЦІОНАЛЬНА СХЕМА МОДУЛЯ ПОПЕРЕДНЬОЇ ОБРОБКИ

ІАЛЦ.467200.003 Д1

Аркушів 1

Київ 2020 р



					ІАЛЦ.467200.003 Д1					
					ФУНКЦІОНАЛЬНА СХЕМА МОДУЛЯ ПОПЕРЕДНЬОЇ ОБРОБКИ					
Зм.	Аркуш	№ Докум.	Підпис	Дата						
	Розробив	Пархомчук О.М.								
	Перевірив	Каплунов А.В.								
	Т. Контр.									
					КПІ ім. Сікорського ФІОТ Група ІО-64					
	Н. контр.	Сімоненко В.П.								
	Затвердив	Каплунов А.В.								
					Аркуш 1		Аркушів 1			

## ДОДАТОК 2

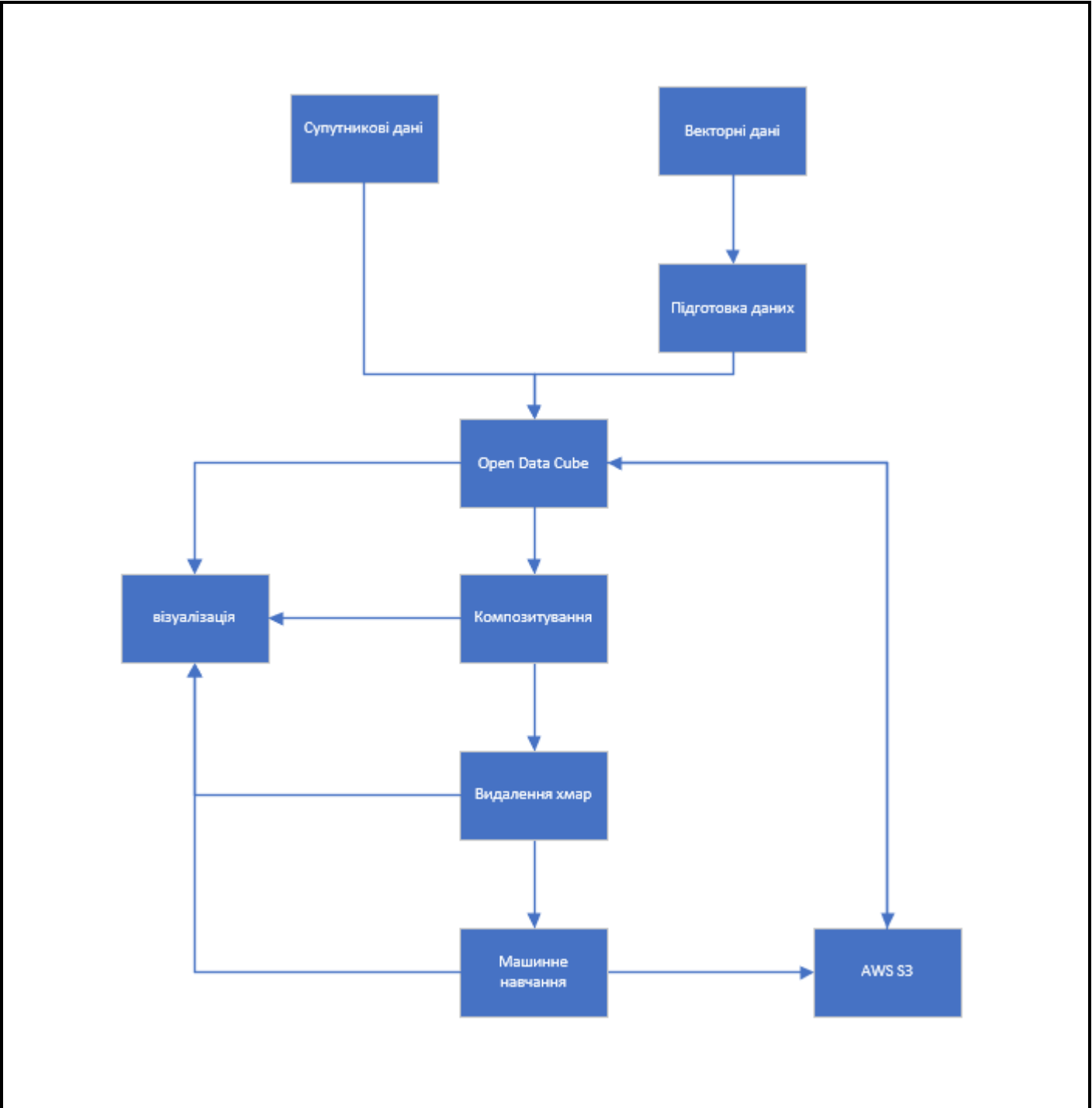
Хмарна система супутникового моніторингу посівів

СТРУКТУРНА СХЕМА

ІАЛЦ.467200.004 Д2

Аркушів 1

Київ 2020 р



					ІАЛЦ.467200.004 Д2										
					СТРУКТУРНА СХЕМА										
Зм.	Аркуш	№ Докум.	Підпис	Дата											
	Розробив	Пархомчук О.М.													
	Перевірів	Каплунов А.В.													
	Т. Контр.														
										Аркуш 1		Аркушів 1			
										КПІ ім. Сікорського ФІОТ Група ІО-64					
	Н. контр.	Сімоненко В.П.													
	Затвердив	Каплунов А.В.													

## ДОДАТОК 3

Хмарна система супутникового моніторингу посівів

ДІАГРАМА КЛАСІВ ДЛЯ МОДУЛЯ ПОПЕРЕДНЬОЇ ОБРОБКИ

ІАЛЦ.467200.005 ДЗ

Аркушів 1

Київ 2020 р





## ДОДАТОК 4

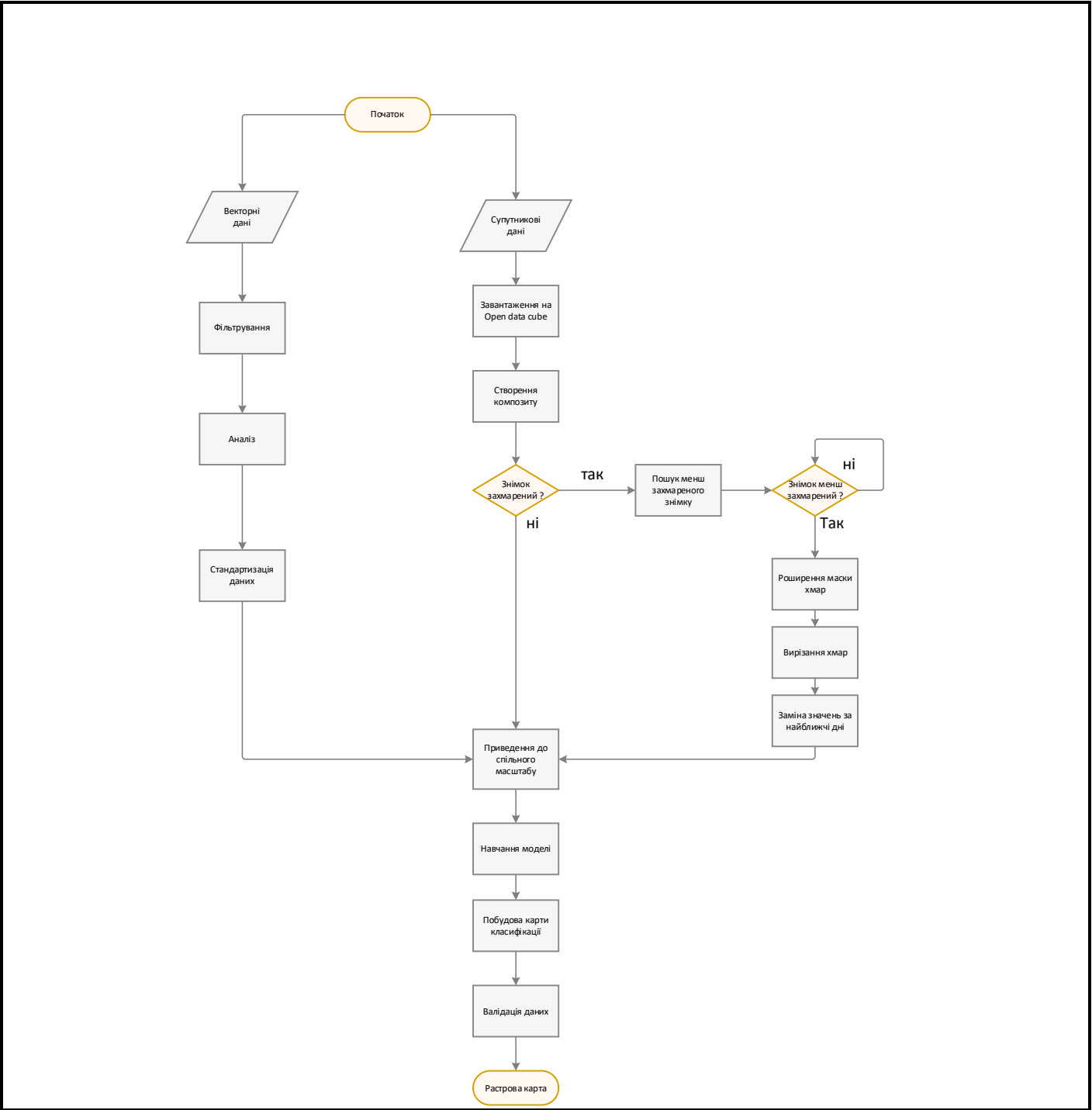
Хмарна система супутникового моніторингу посівів

ПРИНЦИПОВА СХЕМА РОБОТИ

ІАЛЦ.467200.006 Д4

Аркушів 1

Київ 2020 р



					ІАЛЦ.467200.006 Д4					
					ПРИНЦИПОВА СХЕМА РОБОТИ					
Зм.	Аркуш	№ Докум.	Підпис	Дата						
	Розробив	Пархомчук О.М.								
	Перевірив	Каплунов А.В.								
	Т. Контр.									
					КПІ ім. Сікорського ФІОТ Група ІО-64					
	Н. контр.	Сімоненко В.П.								
	Затвердив	Каплунов А.В.								
					Аркуш 1		Аркушів 1			